MAA NOTES #65

Innovations in Teaching Statistics

Joan B. Garfield, Editor

The Mathematical Association of America

Innovations in Teaching Statistics

Library of Congress Catalog Card Number 2005923500

ISBN 978-0-88385-175-3

eISBN 978-1-61444-306-3

Printed in the United States of America

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

Innovations in Teaching Statistics

Edited by

Joan B. Garfield
University of Minnesota

Published and Distributed by
The Mathematical Association of America

The MAA Notes Series, started in 1982, addresses a broad range of topics and themes of interest to all who are involved with undergraduate mathematics. The volumes in this series are readable, informative, and useful, and help the mathematical community keep up with developments of importance to mathematics.

MAA Notes

11. Keys to Improved Instruction by Teaching Assistants and Part-Time Instructors, *Committee on Teaching Assistants and Part-Time Instructors, Bettye Anne Case,* Editor.
13. Reshaping College Mathematics, *Committee on the Undergraduate Program in Mathematics, Lynn A. Steen,*Editor.
14. Mathematical Writing, by *Donald E. Knuth, Tracy Larrabee, and Paul M. Roberts.*
16. Using Writing to Teach Mathematics, *Andrew Sterrett,* Editor.
17. Priming the Calculus Pump: Innovations and Resources, *Committee on Calculus Reform and the First Two Years,* a subcomittee of the Committee on the Undergraduate Program in Mathematics, *Thomas W. Tucker,* Editor.
18. Models for Undergraduate Research in Mathematics, *Lester Senechal,* Editor.
19. Visualization in Teaching and Learning Mathematics, *Committee on Computers in Mathematics Education, Steve Cunningham and Walter S. Zimmermann*, Editors.
20. The Laboratory Approach to Teaching Calculus, *L. Carl Leinbach et al.,* Editors.
21. Perspectives on Contemporary Statistics, *David C. Hoaglin and David S. Moore,* Editors.
22. Heeding the Call for Change: Suggestions for Curricular Action, *Lynn A. Steen,* Editor.
24. Symbolic Computation in Undergraduate Mathematics Education, *Zaven A. Karian,* Editor.
25. The Concept of Function: Aspects of Epistemology and Pedagogy, *Guershon Harel and Ed Dubinsky,* Editors.
26. Statistics for the Twenty-First Century, *Florence and Sheldon Gordon,* Editors.
27. Resources for Calculus Collection, Volume 1: Learning by Discovery: A Lab Manual for Calculus, *Anita E. Solow,* Editor.
28. Resources for Calculus Collection, Volume 2: Calculus Problems for a New Century, *Robert Fraga,* Editor.
29. Resources for Calculus Collection, Volume 3: Applications of Calculus, *Philip Straffin*, Editor.
30. Resources for Calculus Collection, Volume 4: Problems for Student Investigation, *Michael B. Jackson and John R. Ramsay*, Editors.
31. Resources for Calculus Collection, Volume 5: Readings for Calculus, *Underwood Dudley*, Editor.
32. Essays in Humanistic Mathematics, *Alvin White,* Editor.
33. Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results, *James J. Kaput and Ed Dubinsky,* Editors.
34. In Eves' Circles, *Joby Milo Anthony*, Editor.
35. You're the Professor, What Next? Ideas and Resources for Preparing College Teachers, *The Committee on Preparation for College Teaching, Bettye Anne Case,* Editor.
36. Preparing for a New Calculus: Conference Proceedings, *Anita E. Solow,* Editor.
37. A Practical Guide to Cooperative Learning in Collegiate Mathematics, *Nancy L. Hagelgans, Barbara E. Reynolds, SDS, Keith Schwingendorf, Draga Vidakovic, Ed Dubinsky, Mazen Shahin, G. Joseph Wimbish, Jr.*

38. Models That Work: Case Studies in Effective Undergraduate Mathematics Programs, *Alan C. Tucker,* Editor.
39. Calculus: The Dynamics of Change, *CUPM Subcommittee on Calculus Reform and the First Two Years, A. Wayne Roberts,* Editor.
40. Vita Mathematica: Historical Research and Integration with Teaching, *Ronald Calinger,* Editor.
41. Geometry Turned On: Dynamic Software in Learning, Teaching, and Research, *James R. King and Doris Schattschneider,* Editors.
42. Resources for Teaching Linear Algebra, *David Carlson, Charles R. Johnson, David C. Lay, A. Duane Porter, Ann E. Watkins, William Watkins,* Editors.
43. Student Assessment in Calculus: A Report of the NSF Working Group on Assessment in Calculus, *Alan Schoenfeld,* Editor.
44. Readings in Cooperative Learning for Undergraduate Mathematics, *Ed Dubinsky, David Mathews, and Barbara E. Reynolds,* Editors.
45. Confronting the Core Curriculum: Considering Change in the Undergraduate Mathematics Major, *John A. Dossey,* Editor.
46. Women in Mathematics: Scaling the Heights, *Deborah Nolan,* Editor.
47. Exemplary Programs in Introductory College Mathematics: Innovative Programs Using Technology, *Susan Lenker,* Editor.
48. Writing in the Teaching and Learning of Mathematics, *John Meier and Thomas Rishel*.
49. Assessment Practices in Undergraduate Mathematics, *Bonnie Gold,* Editor.
50. Revolutions in Differential Equations: Exploring ODEs with Modern Technology, *Michael J. Kallaher,* Editor.
51. Using History to Teach Mathematics: An International Perspective, *Victor J. Katz,* Editor.
52. Teaching Statistics: Resources for Undergraduate Instructors, *Thomas L. Moore,* Editor.
53. Geometry at Work: Papers in Applied Geometry, *Catherine A. Gorini,* Editor.
54. Teaching First: A Guide for New Mathematicians, *Thomas W. Rishel.*
55. Cooperative Learning in Undergraduate Mathematics: Issues That Matter and Strategies That Work, *Elizabeth C. Rogers, Barbara E. Reynolds, Neil A. Davidson, and Anthony D. Thomas,* Editors.
56. Changing Calculus: A Report on Evaluation Efforts and National Impact from 1988 to 1998, *Susan L. Ganter.*
57. Learning to Teach and Teaching to Learn Mathematics: Resources for Professional Development, *Matthew Delong and Dale Winter*.
58. Fractals, Graphics, and Mathematics Education, *Benoit Mandelbrot and Michael Frame*, Editors.
59. Linear Algebra Gems: Assets for Undergraduate Mathematics, *David Carlson, Charles R. Johnson, David C. Lay, and A. Duane Porter,* Editors.
60. Innovations in Teaching Abstract Algebra, *Allen C. Hibbard and Ellen J. Maycock*, Editors.
61. Changing Core Mathematics, *Chris Arney and Donald Small,* Editors.
62. Achieving Quantitative Literacy: An Urgent Challenge for Higher Education, *Lynn Arthur Steen*.
63. Women Who Love Mathematics: A Sourcebook of Significant Writings, *Miriam Cooney*, Editor.
64. Leading the Mathematical Sciences Department: A Resource for Chairs, *Tina H. Straley, Marcia P. Sward, and Jon W. Scott*, Editors.
65. Innovations in Teaching Statistics, *Joan B. Garfield*, Editor.

MAA Service Center
P. O. Box 91112
Washington, DC 20090-1112
800-331-1MAA FAX: 301-206-9789

Forword

George W. Cobb

The modern introductory statistics course has roots that go back a long way, to early books on statistical methods. R. A. Fisher's *Statistical Methods for Research Workers*, which first appeared in 1925, was aimed at practicing scientists. A dozen years later, the first edition of George Snedecor's *Statistical Methods* presented an expanded version of the same content, but there was a shift in audience. More than Fisher's book, Snedecor's became a textbook used in courses for prospective scientists who were still completing their degrees: statistics was beginning to establish itself as an academic subject, albeit with heavy practical, almost vocational emphasis. By 1961, with the publication of *Probability with Statistical Applications* by Fred Mosteller, Robert Rourke, and George Thomas, statistics had begun to make its way into the broader academic curriculum, but here again, there was a catch: in these early years, statistics had to lean heavily on probability for its legitimacy. During the late 1960s and early 1970s, John Tukey's ideas of exploratory data analysis brought a near-revolutionary pair of changes to the curriculum, first, by freeing certain kinds of data analysis from ties to probability-based models, so that the analysis of data could begin to acquire status as an independent intellectual activity, and second, by introducing a collection of "quick-and-dirty" data tools, so that, for the first time in history, students could analyze real data without having to spend hours chained to a bulky mechanical calculator. Computers would later complete the "data revolution" in the beginning statistics curriculum, but Tukey's EDA provided both the first technical breakthrough and the new ethos that avoided invented examples. 1978 was another watershed year, with the publication of two other influential books, *Statistics*, by David Freedman, Robert Pisani, and Roger Purves, and *Statistics: Concepts and Controversies*, by David S. Moore. I see the publication of these two books 25 years ago as marking the birth of what we regard, for now at least, as the modern introductory statistics curriculum.

The evolution of content has been paralleled by other trends. One of these is a striking and sustained growth in enrollments. Two sets of statistics suffice here: (1) At two-year colleges, according to the Conference Board of the Mathematical Sciences, statistics enrollments have grown from 27% of the size of calculus enrollments in 1970, to 74% of the size calculus enrollments in 2000. (2) The advanced placement exam in statistics was first offered in 1997. There were 7,500 students who took it that first year, more than in the first offering of an AP exam in any subject. The next year more than 15,000 students took the exam, the next year more than 25,000, and the next, 35,000.

Both the changes in course content and the dramatic growth in enrollments are implicated in a third set of changes, a process of democratization that has broadened and diversified the backgrounds, interests, and motivations of those who take the courses. Statistics has gone from being a course taught from a book like Snedecor's, for a narrow group of future scientists in agriculture and biology, to being a family of courses, taught to students at many levels, from high school to post-baccalaureate, with very diverse interests and goals. A teacher in the 1940s, using Snedecor's *Statistical Methods*, could assume that most students were both quantitatively skilled and adequately motivated by their career plans. A teacher of today's beginning statistics courses works with a very different group of students. Most take statistics much earlier in their lives, increasingly often in high school; few are drawn to statistics by immediate practical need; and there is great variety in their levels of quantitative sophistication. As a result, today's teachers face challenges of motivation and exposition far greater than those of a half-century ago.

Fortunately for those of us who teach, the changes I have described have been accompanied, especially over the past quarter century, by an additional set of changes, whose results are represented in this volume. Not only have the "what, why, who, and when" of introductory statistics been changing, but so has the "how." The last 25 years have seen an extraordinary level of activity focused on how students learn statistics, and on how we teachers can be more effective in helping them to learn. One of the remarkable features of this enterprise has been the broad-based and vigorous collaboration between those who come from a background in education research and those who come from a background in statistics. Because of this felicitous collaboration, statistics education has none of the usual fault lines between theory and practice. Instead, theory gets quickly translated into classroom practice, and what goes on in class leads back to a better theoretical understanding.

Until now, the results of this hybrid vigor have been spread mainly by word of mouth, through workshops and papers presented at professional meetings, and through written proceedings and articles in the electronic *Journal of Statistics Education*. One of the leaders in this process has been Joan Garfield, who is known both for her education-based research on the learning of probability and statistics, and as one of a handful of people who have done most to ensure that statisticians and researchers in statistics education work together. Thinking about statistics education from the point of view of the classroom teacher, Joan Garfield recognized that something important was missing. We have a number of good textbooks; we also have a handful of ancillary volumes, on modern content (*Perspectives on Contemporary Statistics*, edited by David Hoaglin and David Moore), on activities for the classroom (*Activity Based Statistics*, by Richard Scheaffer et al.), and on resources (*Teaching Statistics: Resources for Undergraduate Instructors*, edited by Thomas L. Moore). What has been missing is a book of examples of actual classroom practice. Joan Garfield saw the need, marshaled the resources, assembled a team of outstanding teachers, and coordinated their writing. Many of the authors are known for their published work about teaching statistics; many are known from their presentations at workshops and conferences. Now, for the first time, we get a more personal view of what they do in their own classrooms.

Statistics education has come a long way since Fisher and Snedecor. Moreover, teachers of statistics across the country have generally been enthusiastic about adopting modern methods and approaches. Nevertheless, changing the way we teach isn't always easy. In a way, we are all, teachers and learners alike, a bit like hermit crabs: In order to grow, we must first abandon the protective shell of what we are used to, and endure a period of vulnerability until we can settle into a new and larger set of habits and expectations. In this book you can read the stories of many who have done just that.

Contents

1

Introduction

Joan Garfield
University of Minnesota

This is a book of stories about teaching statistics. These stories are told by fourteen different instructors of innovative statistics courses, who demonstrate that learning statistics can be a positive, meaningful, and even exciting experience. Despite the prevailing opinion that statistics courses are dull and difficult for students, these stories paint quite a different picture. In the classes of the instructors whose stories fill this book, students are engaged in learning, are empowered to do statistics, and appreciate the instructional methods of their teachers. The instructors profiled in this book are inspiring, dedicated teachers who have devoted considerable effort to creating courses and materials that enable students to successfully learn statistics.

Many of these stories were initially collected as part of a grant from the National Science Foundation. In 1998 I began a two-year study of teachers of first courses in statistics. My main goal was to determine how recommendations for educational reform were having an impact on the teaching of these courses. During the 1970s and 1980s, much dissatisfaction had been expressed about first courses in statistics. Students were unhappy with these classes, instructors were dissatisfied with the students and their inability to learn and apply course content, and there was a general perception that most introductory courses were ineffective and unpleasant. In response to the numerous complaints, various groups began to address the need to change, or reform, the introductory statistics course.

One group that issued an influential report was a joint committee of the Mathematical Association of America (MAA) and the American Statistical Association (ASA). Headed by George Cobb, this group produced a report in the MAA Notes volume, *Heeding the Call for Change* (Cobb, 1992). This report urged instructors to emphasize statistical thinking by including more real data, and by using fewer formulas, to use more activities and have fewer lectures. Back in 1998, six years after this report was published, I wondered how many faculty teaching statistics were aware of basic reform recommendations and were using them to restructure their courses.

I also wondered how many faculty had heard and tried to respond to exhortations by David Moore (1997a), regarding changes in course content, pedagogy, and uses of technology. Although there seemed to be quite a bit of agreement among statisticians about teaching the introductory course (see Scheaffer, 1997), I wondered how many non-statisticians were aware of the recommendations and were in favor of reform efforts. As

Moore (1997b) has pointed out, many people teaching introductory statistics are not statisticians and far more sections of introductory statistics are taught in mathematics departments or in other disciplines than by statisticians in statistics departments.

In response to the many papers and presentations on the need for educational reform, numerous projects were funded by the National Science Foundation and other organizations to implement aspects of this reform (see Cobb, 1993). I was curious as to how these projects, along with writings, presentations and workshops, were being viewed by teachers of statistics, and how they were bringing about needed changes. I received a small grant from the National Science Foundation to conduct this study. One part of the project was developing a six-page survey, which was administered by paper, email, or via a web format to teachers of introductory statistics courses. The results have been summarized in Garfield, Hogg, Schau and Whittinghill (2002), and a complete report is available online at

www.coled.umn.edu/EdPsych/Projects/Impact.html.

Another component of my project was a series of case studies of exemplary instructors who had successfully implemented reform recommendations in their introductory statistics courses. These case studies were used to describe the process of changing one's course and to provide a more detailed picture of what some "reformed" courses look like. A small group of teachers were interviewed who were known to be teaching innovative courses or whose responses on the survey suggested that they were incorporating reform recommendations. These individuals came from different departments and types of institutions. They were asked to describe the key features of their introductory course, how their course differs from a "traditional" course, the process that led them to develop their course, what types of support they received, and how the course will continue to be revised in the future.

The results revealed surprising differences from course to course and illustrate the complexities of teaching in different institutions and departments. Although all instructors were implementing some reform recommendations, the nature and extent of the implementation varied, sometimes due to available resources, sometimes due to the characteristics of students at a particular institution, and often due to the instructor's experience and beliefs about teaching.

For example, when asked how their course differs from a traditional course, the responses included:

- I teach statistics as a language course, and try to help the students develop literacy about statistics.
- I have students keep journals of both statistical problems and reactions to the course.
- There is no memorization required of students. On exams, I give credit for effort and explanation.
- I use a mastery exam (scored but not graded), which students must pass, like a drivers' test, before they are allowed to carry out a real statistical investigation.
- I use lots of pairs and group work.
- I emphasize data production and simulation.
- Students have many opportunities for self-assessment.
- I create an interactive learning environment.
- I use two types of technology tools in my class; Minitab for Homework and projects, Fathom for illustrating and developing concepts.
- I use the PACE model to create a highly interactive learning-centered classroom. PACE stands for **P**rojects, **A**ctivities, **C**ooperative learning in a computer-based classroom environment, and reinforcement through **E**xercises.

Despite the differences listed above, there was also a common theme among many instructors who stated that they focus more on concepts and big ideas and on data analysis and interpretation and less on computation, formulas, and theory.

The process that led these instructors to their current course often included conversations with other statistics educators, reading articles in the *Journal of Statistics Education* or listening to presentations at professional meetings, and trial and error testing of new techniques. Challenges faced along the way included lower teaching evaluations due to problems that arise when trying new techniques for the first time, the lack of rewards for effort applied to teaching (as opposed to research), student resistance to changing from passive to active learners (where more is demanded from them), and colleagues who want to see introductory statistics courses with more math, in particular, probability and more rigor.

The interviews revealed that these instructors had spent a great deal of effort thinking about their courses, and had dedicated huge amounts of time to improving and revising their courses. Although generally pleased with the results, most shared ideas they had for how they will continue to make changes and indicated that their courses are still being developed. Some felt that they are "moving in the right direction" but still "have a ways to go." Some reported that each time they teach it's a different course. Others commented that the first time they taught a "reform course" was difficult but that things went better the next time.

One instructor commented on the changing population of students who work more hours at full-time or part-time jobs, do not read newspapers, and have less interest and motivation. In order to find topics that interest her students she reports being "pretty much down to weather, cell phones, and fast food." Another remarked that "students struggled in the course but many learned a lot and were able to retain a fair bit. Often they didn't appreciate that they were learning more until they saw how much other students struggled in later courses. Students began to appreciate the prevalence of statistics in everyday life, and how much more cautious we should be using statistical statements and interpretations". Some instructors were pleased to see much better quality in student projects that were well written and used appropriate graphs and analyses. Others noted increases in student satisfaction and attitudes about statistics. One instructor commented that "a large majority of my students now see this course as a positive experience."

A number of instructors indicated that they have been able to devote their time and effort to teaching because of having tenure and academic freedom. Some enjoyed freedom to experiment with their course because no one in their department knows or cares about how they are teaching the courses (one instructor referred to this as "benign neglect"). A few instructors appreciated the support of a department chair or colleagues or have received internal or external funding to support their efforts. A consistent result was that most of the faculty studied cited colleagues from outside of their institution as their main source of teaching support, particularly those they see at their professional meetings.

After finishing this project and presenting the results at several different conferences, it was suggested that I use these case studies as the basis of an edited book, and that is what led to the production of this volume. A few of the original faculty interviewed chose not to write a chapter for this book, so some substitutions were made.

Each chapter is written by a faculty member who is teaching an innovative course in statistics. The book is organized into four sections. The first section consists of statisticians who teach in departments of mathematics. Section 2 consists of stories by faculty who work in community colleges or with nontraditional, under-prepared students. The third section is written by statisticians who teach in departments of statistics, and the fourth section tells the stories of statistics instructors who work with students in other areas, specifically business and psychology.

The fourteen innovative teachers of statistics in this book each tell their own story about how they teach the introductory course. They begin by describing how they became a teacher of statistics, which was typically not part of their initial career plans. They provide details about the course they teach, describing their teaching method, textbook, types of student assessments, and uses of technology. One typical class is described in detail, to provide a snapshot of what each person's teaching looks like. The writers then tell the story of the process they went through in developing an innovative course, and conclude the chapter with future plans. Except for Jonathan Cryer, who was about to retire from teaching, all have plans for how they plan to continue to change and improve their courses. As you read these stories, you will learn about some great activities, some helpful technological tools and some innovative assessment methods.

I hope that by reading these stories, teachers of statistics will understand and be motivated to try different ways to implement reform recommendations, so that these recommendations may continue to lead to a variety of new materials, activities, and teaching approaches. I hope that teachers of statistics may be inspired by these stories and feel encouraged to try new methods, leave behind more traditional approaches to teaching statistics, and carefully examine the effect of their teaching on student learning.

I hope that this book will be helpful to graduate students preparing to teach statistics at the college or high school level. By providing models and illustrations of exemplary teaching, future statistics teachers may be encouraged to implement and build on these approaches. If all these things happen, the ultimate benefit will be experienced by students as they enroll in improved statistics courses and have positive experiences learning, doing, and using statistics.

References

Cobb, G. (1992). Teaching statistics. In L. Steen, (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 3–43). Washington: Mathematical Association of America.

Cobb, G. (1993), Reconsidering Statistics Education: A National Science Foundation Conference, *Journal of Statistics Education* [Online], 1(1) www.amstat.org/publications/jse/v1n1/cobb.html.

Garfield, J., Hogg, B., Schau, C., and Whittinghill, D. (2002) "First Courses in Statistical Science: The Status of Educational Reform Efforts" *Journal of Statistics Education* [Online], 10(2) www.amstat.org/publications/jse/v10n2/garfield.html.

Moore, D.S. (1997a). New pedagogy and new content: the case of statistics. *International Statistical Review*, 65, 123–137.

Moore, D.S. (1997b). Response. *International Statistical Review*, 65, 162–165.

Scheaffer, R.L. (1997). Discussion. *International Statistical Review*, 65, 156–158.

Section 1

Statistics Teachers in Mathematics Departments

2

Teaching a Data-Oriented, Activity-Based Course

Allan J. Rossman
California Polytechnic State University, San Luis Obispo

I have no regrets at all about my career choice or about concentrating my scholarship on statistics education. The dedication and camaraderie among statistics educators has been inspiring, and working together to try to improve the teaching of statistics is quite gratifying.

Becoming a teacher of statistics

I knew prior to attending graduate school in statistics at Carnegie Mellon University that I wanted to become a teacher of statistics at the undergraduate level. My interest in statistics had begun at an early age in a gender-stereotypical fashion: I was fascinated by sports statistics and loved to play table-top board games that simulated sports seasons and allowed me to tabulate statistics to my heart's content. Fortunately, I became even more interested in statistics upon learning that it was really a scientific discipline about gaining insights from data and not just a hard-to-pronounce synonym for earned run averages and field goal percentages. My interest in teaching was nurtured by my undergraduate training at Geneva College, a small liberal arts college where instructors' primary responsibility was teaching. My instructors taught a wide variety of courses, devoted considerable energy and creativity to their teaching, and took great pride in their students' intellectual development and achievements. I received a taste of the teaching profession by serving as a grader, tutoring session leader, and occasional substitute teacher, and I decided that I wanted to devote my professional life to undergraduate teaching.

My career goal was quite unusual among my fellow graduate students at Carnegie Mellon, the vast majority of whom were preparing to pursue careers at research institutions or in industry. To prepare for this goal I served as a teaching assistant for a large introductory course, and I taught courses on my own during summers. I also attended lectures and workshops held by the university's Teaching Center. I was very pleased to be hired by Dickinson College in May of 1989, and I finished my PhD in Statistics that summer before beginning my teaching career in the fall. Dickinson is a private liberal arts college of about 2000 students in Carlisle, Pennsylvania.

During the summer of 1989 I read an article that profoundly affected my career. It was written by Tom Moore and Rosemary Roberts and was titled "Statistics at Liberal Arts Colleges" (Moore and Roberts, 1989). I was delighted to read that there were other statisticians who had chosen to pursue teaching-oriented careers at liberal arts colleges, and I was especially heartened to read that they enjoyed the challenges and opportunities provided by that career choice. Additionally, this article introduced me to the term "data-driven," as it argued that statistics courses should be redesigned to use real data in focusing on issues associated with data collection, analysis, and interpretation. Moreover, the article asserted that faculty at liberal arts colleges, with their emphasis on teaching and their freedom to experiment pedagogically, are especially well positioned to lead this effort. I felt quite inspired to accept this challenge and play a role in the reform of statistics education.

After arriving at Dickinson College I began to familiarize myself with literature from the mathematics community. I was strongly influenced by David Moore's article in *The College Mathematics Journal*, which he titled "Should Mathematicians Teach Statistics?" and began with a resounding "No!" (Moore, 1988). This article helped me to understand some of the differences between the disciplines of statistics and mathematics and the importance of teaching statistics for its own sake rather than as a sub-discipline of mathematics.

An important event in my development as a statistics teacher and my decision to concentrate on statistics education was the 1992 Winter Meeting of the American Statistical Association, which was held in Louisville and focused on the theme of education. I learned a great deal about projects and ideas that educators across the country were involved in, and I was pleased to see so many statisticians deeply concerned with educational issues. I was also gratified to see that the ideas I espoused were well received and not out of step with others' views. A highlight of the meeting was the keynote address by Joan Garfield, which addressed the topic of how students learn statistics and implications of the theory of constructivism for teaching statistics. Joan later published an article based on this address in the *International Statistical Review* (Garfield, 1995). It was very helpful for me to learn more about the theoretical background behind the "self-discovery" pedagogy that I was developing.

Another important influence around this same time was the publication of the recommendations for teaching introductory statistics from the ASA/MAA Joint Committee on Undergraduate Statistics in 1992 (Cobb, 1992). In a nutshell these suggestions were to teach statistical thinking, emphasize more data and concepts, and foster active learning. I was very encouraged that the recommendations by this national group were consistent with my approach.

I have no regrets at all about my career choice or about concentrating my scholarship on statistics education. The dedication and camaraderie among statistics educators has been inspiring, and working together to try to improve the teaching of statistics is quite gratifying.

My course

I spent the first twelve years of my teaching career at Dickinson College, before moving to California Polytechnic State University in 2001. As the sole statistician in Dickinson's Department of Mathematics and Computer Science, one of my primary responsibilities was to teach the course in Elementary Statistics (Math 121). This course was required of Economics majors and could be used as a major elective in fields such as Policy Studies and Environmental Studies. The biggest audience for the course, though, was students satisfying a general education distribution requirement. These students were typically majoring in humanities and social sciences, and many were following a pre-medicine program.

One of the very best things about my position was that I was encouraged, and given the freedom, to re-design the course as I deemed appropriate. The course that I inherited used a very traditional textbook and focused on students performing calculations by hand. Typical of the course was the following exam item: Students were presented with five numbers and were asked to calculate a confidence interval for the mean. No context was provided, and no definition of μ was given; no conceptual understanding was expected, and no interpretation of results was required.

A first step in redesigning the course was to adopt a new textbook. The first edition of Moore and McCabe's *Introduction to the Practice of Statistics* (Moore and McCabe, 1988) had just been published, and I considered this book with high expectations. I had taught from Moore's previous text, *Statistics: Concepts and Controversies* (Moore, 1985) in graduate school and had not only enjoyed but also learned a lot from this unconventional text. I was pleased to find the new book by Moore and McCabe had some of the unique features of Moore's earlier book but was more mainstream and more appropriate for my course. This change in textbooks allowed me to shift the focus of the course to data analysis and to put more emphasis on understanding concepts and on interpreting results. This move alarmed some of my mathematician colleagues who also taught the course, as they found Moore & McCabe to be much less mathematical than they were comfortable with.

I also began to integrate Minitab assignments into the course. Dickinson was running Minitab on a mainframe at that point, but statistics students had not been expected to do much with it. I began asking students to complete assignments that involved using Minitab to analyze data and to simulate long-run behavior of random phenomena.

After adjusting the content of the course to focus on data and concepts, I began to change the pedagogy of the course. Led by my colleagues Nancy Baxter Hastings (in Mathematics) and Priscilla Laws (in Physics), I received funding from NSF to equip computer classroom laboratories and from FIPSE to support curriculum development. I started this transition by creating worksheets of activities for students to complete during class time while reading Moore & McCabe outside of class. As the semesters progressed, I expanded and refined the activities, placing less emphasis on the text and putting more exposition on the other materials. As I observed students' reactions to the materials and assessed their learning from them, I gradually came to depart from Moore & McCabe in a few areas such as sequencing of topics. Eventually I arrived at the point where the text was dropped and my materials covered the entire course. These curricular materials became *Workshop Statistics* (Rossman, 1996), a collection of activities that guide students to discover statistical concepts, explore statistical principles, and apply statistical techniques.

The course that evolved from this process had as its primary goals to help students develop:

- the ability to apply and interpret the results of a variety of statistical techniques, including both exploratory and inferential methods;
- an understanding of many of the fundamental ideas of statistics, such as variability, distribution, association, causation, sampling, experimentation, confidence, and significance;
- a critical perspective with which to analyze and assess statistical arguments such as one encounters in the popular press as well as in scholarly publications.

In terms of content, there was more of a focus on data analysis and much less emphasis on probability than in most traditional courses. The content was similar to Moore & McCabe and to the current Advanced Placement Statistics Course (College Board, 2002), although my course covered much less probability and fewer techniques than those two courses. Technology was used in virtually every class period and homework assignment, with Minitab being the software employed.

A typical class

I enjoyed the luxury of teaching every class session in a computer-equipped classroom/lab where the students sat in pairs at each computer. The lab tables also had lots of room for working together on activities that did not involve computers. The class was limited to 24 students, and it typically met for 75-minute periods on a Tuesday/Thursday schedule.

On a typical day, I would start by asking for questions on the previous day's topic and assignment. Then I would give a brief introduction to the day's topic and where it fits in with what has come earlier. Often we would collect some class data to be analyzed that day or perhaps do a simulation together and pool results. Then I would ask students to work through the activities in their book.

The majority of class time was spent with students working in pairs on activities designed to lead them to "discover" statistical ideas, to explore statistical principles, and to analyze data. I saw my primary role as instructor being to facilitate that learning. This facilitation took many forms, but I spent most of the class time circulating about the classroom, looking over shoulders, answering students' questions, questioning them about their responses to the activities, and pointing students in the right direction. I made a concerted effort to be "proactive" in the sense of engaging students in conversation about the material and activities without waiting for them to ask questions of me. I tried to visit every pair of students several times per class session to make sure that they were on task and on track.

Providing timely and helpful feedback is very important with this approach. In addition to providing feedback as I moved about the room, I would also write answers on the board (for those questions with definitive right/wrong answers) and ask that students check their work often. Whenever a common misunderstanding arose, I would stop the class and deliver a "mini-lecture" to try to clear up confusion quickly. I tried only to lecture on a topic after students had wrestled with it through an activity, the hope being that they would better understand what I had to say if they had already thought about the issues. At the end of the class session, I would devote several minutes to summarizing the key points of the day, usually by eliciting student responses to my probing questions about the activities.

Some sample activities

Let me illustrate by describing some activities through which my students study the topics of association and correlation. After an introduction to the definition of association and the notion of direction and strength, students consider nine scatterplots of various pairs of variables involving car models, such as miles per gallon rating versus weight. Students work together on the task of arranging the scatterplots in order from that with the strongest negative association through virtually no association to the strongest positive association. Before students see the formula for the correlation coefficient or any of its properties, they return to the same data and use technology to calculate the correlation coefficient for the nine pairs of variables represented in those scatterplots. They find that the correlation values increase monotonically with their ordering of the scatterplots. Then students answer a series of questions that lead them to discover basic properties of the correlation coefficient: that it has to be between –1 and +1 (inclusive), that the sign of the correlation coefficient reflects the direction of the association, and that a correlation closer to +1 or –1 indicates a stronger (linear) association between the variables. Students next analyze data on prices and rents of Monopoly game properties, and they are asked to use technology to investigate the effect on the correlation coefficient of changing one value. They discover that the correlation coefficient is not resistant to the effects of influential observations. The next activity asks students to analyze the relationship between the life expectancy of a country and the number of people per television set in the country, and they find a fairly strong negative association. When asked whether sending televisions to less wealthy countries will therefore cause their inhabitants to live longer, students recognize the folly of that argument and realize that association does not imply causation.

With all of these activities students work through the analysis and draw conclusions working in groups of peers. My role as instructor is to prod them in the right direction, and at the end of the class period I offer a summary of what they were to have learned. My hope, though, is that students develop a deeper understanding of these concepts by working through them rather than being told them.

Student assessment

Assessment methods used in my course were fairly traditional in that I relied on in-class exams and homework assignments to determine grades. I assigned 5–7 homework activities per week and gave three exams over a semester. I tried to write exam and homework questions that concentrated heavily on conceptual understanding, on interpretations of results, and on communications skills. My goal was to write exams and assignments so that at least 50% of the points pertain to conceptual/ interpretive ability rather than computational/ mechanical skills. The exams were open-book and open-notes, partly to emphasize the conceptual nature of the questions and partly to encourage students to write good explanations on their in-class activities.

I would have liked to include a computer component on the exams, but with half as many computers as students this was not feasible. As a compromise I would provide students with computer output, including irrelevant output at times, on exam questions, asking them to identify and interpret the relevant portions.

While my curriculum materials were in early stages of development, I collected and graded students' in-class activities as well as homework activities. This practice was invaluable for helping me to see which activities were working well and which were not, but it was too time-consuming to continue indefinitely.

Toward the end of my tenure at Dickinson I began to incorporate "mini-projects" into the assessment as well. These assignments asked students to tackle slightly more open-ended questions that involved larger data sets and use of technology and to write a report of their analyses and findings. One example, shortly after the 2000 election, was asking students to analyze election data for a given state. For part of this project they analyzed county-by-county election returns and developed a model for predicting the Presidential vote in the county based on the votes for another statewide office. For the other part of this project students analyzed exit poll results and carried out inference procedures for generalizing the exit poll findings to the population of voters in the state.

I did not perform rigorous evaluations of my course's effectiveness, but colleagues in other departments, principally in social sciences, were pleased with the emphasis on data and concepts and were satisfied with the knowledge that students brought from my course to their Research Methods courses. I was fortunate that colleagues in "service" departments were very collegial and happy to discuss the teaching of introductory statistics; we wrote an article summarizing these conversations (Chromiak et al, 1992). I personally felt that students were learning the "big ideas" of statistics and were better-informed consumers of statistical information than they had been with the previous version of the course.

Developing an innovative statistics course

My colleagues at Dickinson were very influential in my early career and on my decision to pursue statistics education as the focus of my scholarly endeavors. My colleague Nancy Baxter Hastings was pursuing curriculum development and pedagogical research in mathematics education, and another colleague Priscilla Laws was developing "Workshop Physics," a project that replaces the distinction between lecture and laboratory in introductory physics courses with an entire curriculum designed around student investigation and discovery. When I learned of Nancy's and Priscilla's ideas concerning constructivism and active learning, my immediate reaction was that this pedagogy is ideally suited to the study of statistics—what better way to learn the science of data than by engaging in the analysis of data?

I was very fortunate that Dickinson supported and encouraged innovative work in pedagogy and curriculum development. I greatly benefited not only from having colleagues with similar interests to talk with but also from Dickinson's personnel evaluation process giving "credit" for this type of scholarship as long as it is favorably peer-reviewed.

I also found it very helpful to look beyond my own institution by reading articles about others' projects and attending conferences where I could see their presentations first-hand. I mentioned earlier some of the articles and presentations that had the greatest impact on my own teaching. I learned early on that it is important not to work in isolation. As the only statistician at Dickinson, I was very fortunate to become involved with the network of "isolated statisticians" who meet at conferences and hold discussions via e-mail about a myriad of teaching issues. My own teaching and scholarship have been enhanced greatly by the infusion of fresh perspectives from collaborators, particularly Beth Chance who joined me in writing a substantial revision for the second edition of *Workshop Statistics* (Rossman and Chance, 2001). One practice that has been especially helpful, and which I highly recommend to others, is to always ask a colleague (in my case, Beth Chance) to critique exam questions and other assessment items. This invariably leads to my asking better and sharper questions of students and gaining better information about what they know and where they need help.

Future plans

One of my current projects is to extend the active-learning, data-oriented approach to statistics courses for mathematically inclined students. Most of the reform efforts in statistics education have focused on the introductory service course, which to some extent has left mathematics and statistics majors behind. This problem is especially important for prospective teachers who need to experience the type of course that they will be expected to teach, a point emphasized in the Mathematical Education of Teachers document from the Conference Board on Mathematical Sciences (CBMS, 2001). My colleague Beth Chance and I have an NSF grant to develop curricular materials that maintain a focus on data and a spirit of active learning while taking advantage of the mathematical backgrounds and abilities of this student audience (Rossman and Chance, 2002; Chance and Rossman, 2006).

My current challenges are to adjust to teaching my course under the quarter system at my new institution and to teaching a broader spectrum of introductory courses for different student majors. One specific goal is to make better use of the new software products now available that have been designed to facilitate student explorations of concepts. I will soon be using Fathom Dynamic Statistics software as the primary technology in my course for the first time. I also hope to learn more about course management software and use it to facilitate communication with and among my students.

Of course, one of the things about our profession that is simultaneously exhilarating and infuriating is that one never masters it. In many ways my future plans are exactly the same as they have been since my first year of teaching: trying to figure out new ways to engage students and help them with their study of statistics. I hope that I never fool myself into believing that I've discovered the secret and no longer need to innovate.

References

Chance, B., and Rossman A. (2006). *Investigating Statistical Concepts, Applications, and Methods*. Pacific Grove, CA: Duxbury Press.

Chromiak, W., Hoefler, J., Rossman, A., and Tesman, B. (1992). A multidisciplinary conversation on the first course in statistics. In F. Gordon and S. Gordon, (Eds.), *Statistics for the Twenty-First Century* (pp. 26–36). Washington, DC: Mathematical Association of America.

Cobb, G. (1992). Teaching statistics. In L. Steen, (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 3–43). Washington: Mathematical Association of America.

College Board (2002). AP statistics course description. New York: The College Board. www.collegeboard.com/ap/statistics/.

CBMS (2001). *The mathematical education of teachers*. Washington: Conference Board of the Mathematical Sciences.

Garfield, J. (1995). How students learn statistics. *International Statistical Review*, 63(1), 25–34.

Moore, D. (1988). Should mathematicians teach statistics? (with discussion). *The College Mathematics Journal,* 19, 3–35.

Moore, D. (1985). *Statistics: Concepts and controversies* (2nd ed.). New York: W.H. Freeman.

Moore, D., and McCabe, G. (1988). *Introduction to the practice of statistics* (1st ed.). New York: W.H. Freeman.

Moore, T., and Roberts, R. (1989). Statistics in liberal arts colleges. *The American Statistician*, 40, 80–85.

Rossman, A. (1996). *Workshop statistics: Discovery with data* (1st ed.). New York: Springer-Verlag.

Rossman, A., and Chance, B. (2001). *Workshop statistics: Discovery with data* (2nd ed.). Emeryville, CA: Key College Publishing.

——— (2002). A data-oriented, active learning, post-calculus introduction to statistical concepts, methods, and theory. In B. Phillips, (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*, Cape Town, South Africa.

3

Using the PACE Strategy to Teach Statistics

Carl Lee
Central Michigan University

Life is unpredictable. However, when I look back on the path I took to become a statistics teacher, it seems it had been laid out for me and all I did was walk along this 'pre-determined' path.

Becoming a teacher of statistics

My journey to become a statistician and to teach statistics in a university is perhaps quite different from most statistics educators. My undergraduate degree was in Agronomy from the National Taiwan University in 1976. At that time, there were no statistics programs nor any statistical software packages available in Taiwan. One could take courses in either "Probability & Measure Theory" in the Mathematics Department or "Biometry" in Agricultural Sciences. My statistics and computer programming training took place primarily in the Biometry program, within the Department of Agronomy.

One reason that I was attracted to the study of statistics was the influence of a particular professor on my life. When I was a sophomore in college, Professor Chang asked me to join his experimental design laboratory as a student research assistant. In this job I had many opportunities to help conduct research studies and to help develop computer programs to analyze data. This experience was what motivated me to want to pursue the study of statistics.

After receiving a degree in Agronomy, my first full-time job was as a computer programmer working at the computer center at National Taiwan University. I was responsible for developing statistical software and I also served as a statistical data analyst. However, what I really wanted to do was to be a university professor. A professor's job was very respectable in Taiwan and paid well. I thought I would enjoy being a professor of either Computer Science or Statistics. In order to become a professor, I first needed to earn a doctoral degree. I decided to first pursue a Master's degree in Mathematics, an area that seemed essential for doctoral work in either Computer Science or Statistics. I enrolled in a Master's in Mathematics program at the University of West Florida and was later accepted into the PhD program in the Department of Statistics at Iowa State University (ISU).

I found that the diversity of courses and opportunities at ISU fit my interests very well. I took as many courses as possible and continued to dream of becoming a university professor. Professor Kempthorne's influence was perhaps the most important in my five years at ISU. He trained me to be an independent researcher and inspired me with his passion for teaching and research. I started my journey in agricultural science in Taiwan, and ended up teaching statistics in the Department of Mathematics at Central Michigan University (CMU).

When I first taught the introductory statistics course in the fall of 1984 at CMU, I looked out at 90 strange faces watching me. I was very nervous, and I worried about my imperfect English skills. I have to admit that I had a lot of nightmares about teaching that first year. At the end of the semester, the student evaluation results were mixed, containing some nice comments and many unflattering ones. I took these comments very seriously and reviewed my teaching experience, looking for ways to improve my course instruction.

I knew I was well prepared, excited about the subject, and was willing to provide extra help to students. My applied science and statistical consulting background also helped me to provide good examples and to keep the class activities more interesting. Occasionally, I provided students with information on Asian culture and Chinese characters to engage them and stimulate their interest. I even prepared some special jokes for the class to make the students laugh and feel more comfortable. I really thought I did my best and did it well. However, my results from the student course evaluations were below the departmental average, and that was very disappointing to me. As a junior faculty, I had been told to focus on publishing articles but to make sure that my teaching evaluations were as good as the departmental average. I was not even close!

I tried to understand what was wrong with my teaching. I had a sense of urgency in wanting to improve my teaching to at least the departmental average. The first thing I did was to request smaller class sizes. In order to do this, I offered to teach more sections. Building on student comments and suggestions, I began to prepare additional handouts for difficult concepts. I tried to show more confidence by speaking with a louder voice, made more eye contact and attempted to have more interaction with students by calling their names. I conducted help sessions before tests. I also introduced a simple five-minute midterm teaching evaluation, giving students the opportunity to comment on my teaching and offer suggestions that I could try to implement during the second half of the semester. I also prepared even more jokes to liven up my class. I began to introduce student projects and use more technology in the classroom.

Over a period of ten years the student course evaluations showed steady and impressive improvements. I was even nominated for teaching awards, which made me feel very appreciated for my efforts to improve my teaching. I received tenure and promotion, and finally felt comfortable with my teaching. Class size was no longer a problem. I taught classes with sizes ranging from 30 to 150 students. During this time period, I was also in charge of the university statistical consulting service, which I enjoyed greatly. My teaching career turned out to be pretty well after all!

In 1995 the Department of Mathematics introduced a PhD program with an emphasis in college teaching. This PhD program has a unique goal of training students to be excellent teachers of Mathematics at the college level. Although the Department of Mathematics has a strong group of Mathematics Educators emphasizing K–12 teaching, there was an immediate need for other faculty to become involved in studying and taking on the scholarship of teaching and learning at the college level. Remembering the powerful influence on my own life due to studying and working with Professors Chang and Kempthorne, I decided that I wanted to participate in this new program.

I began to rethink my teaching from the perspective of how much students actually learn in my classes. Reading articles related to student learning and new teaching pedagogy became a major activity of mine. During this time period, I began to pay increased attention to student learning outcomes. As part of my transformation, I became involved with the university-wide Student Learning Outcomes Assessment in 1997 and became the University Assessment Coordinator from 1999 to 2001. This was an exciting learning experience for me and has greatly influenced my own teaching and assessment practices.

Over the past seven years, I have spent a lot of time designing an alternative teaching approach for my introductory statistics courses. I call this approach the "PACE" model. PACE stands for **P**rojects, hands-on **A**ctivities, **C**ooperative learning in a computer classroom environment, and **E**xercises for re-enforcement. I developed the PACE model as a result of wanting to find out how students best learn quantitative skills and how different teaching pedagogies impact students who have little motivation to learn statistics.

The PACE model is based on the following principles (see Garfield,1993):

- People learn better by constructing knowledge themselves through guided processes.
- Active problem solving through teamwork promotes active learning.
- Practice and feedback are essential ingredients for understanding new concepts.

The PACE strategy provides a structured approach that integrates projects and hands-on activities, which are conducted cooperatively, in a computer environment. This approach places students at the center of the learning process, where the instructor acts as a facilitator who leads and guides students to discover and understand new concepts. Statistics is approached as a scientific tool for solving real-world problems. Students have many opportunities to work as a team, and are engaged in writing reports and giving oral presentations.

My courses

Central Michigan University (CMU) is located in Mt. Pleasant, near the center of the lower peninsula of Michigan. CMU is a state-funded regional research university. It has an on-campus student population of about 19,500 and the extended learning college has an additional 9,000 students. More than 70 percent of the on-campus students are from within a radius of 200 miles of the university. The student population is predominately Caucasian (approximately 93 percent) with about 45 percent male and 55 percent female. The on-campus students range in age from 18 to 25; there are very few adult learners.

There are two introductory statistics courses at CMU. One is for students in the College of Science & Technology, and it is called Elementary Statistical Analysis, STA 382. The other course is for students in other colleges, and it is called Introduction to Statistics, STA 282. Approximately 75 percent of the students in STA 282 are from the College of Business. Therefore, this course emphasizes business applications, and the textbook has been selected to reflect the appropriate applications in business

Students who take STA 282 often have little mathematics background. The prerequisite for the course is college algebra. At the beginning of my own section of this course I give a diagnostic test that measures basic algebra skills and concepts. The test results usually follow an approximately uniform distribution. A similar test is given in STA 382 which has as a prerequisite pre-calculus. The results are much better than for STA282, although surprisingly, there are always a few students who have difficulty manipulating fractions.

Generally, STA 382 students are better disciplined and have better attitudes toward quantitative subjects than do students in STA 282. In terms of differences in content, both of these courses cover the same basic topics offered in an introductory statistics course, but STA 382 covers additional subject matter including Poisson and hypergeometric distributions, normal approximation to binomial distribution, and sample size determination.

My primary learning goals for students in both courses include being able to:

- read and distinguish appropriate data analysis from inappropriate data analysis in the media,
- carry out a proper process of collecting data, describing data, and summarize and present the results to an audience that does not have a strong quantitative background,
- identify and compare variability in data in different situations under different contexts,
- appreciate the value of and apply statistical reasoning skills to their own discipline(s).

I also try to develop students' statistical thinking, as described by Wild and Pfannkuch (1999). I use the PACE model to help achieve these learning goals. Each component of the PACE model is described below (see also Lee, 2000).

Projects provide students opportunities to go through the entire process of data collection, data production, data summarization, data analysis, report preparation, and result presentation. Structured, self-selected projects seem to produce more student motivation and interest. The concepts of sampling design, experimental design, independent two-sample data, paired sample data, categorical data, regression, and correlation data are introduced very early on. Students are then instructed to look for projects that feature the properties of independent two-sample data, paired sample data, and regression and correlation data respectively, using any sources that are available to them. The commonly used resources are sports data, weather data, and studies related to the students' own areas of expertise.

Activities are the core component of each class (see stat.cst.cmich.edu/statact/). Each activity is designed to introduce several new concepts and to review previously learned concepts. Students are guided to work in pairs to conduct the activity, collect data, analyze the data, and present their findings. For example, the activity 'How far are you from home?' is the first activity conducted on the first day of class. This activity is designed to help students discover the relationship among population, samples, selection of measurement, random sampling, graphical presentation, mean, median, standard deviation, percentile, and shape of distribution. Many of these concepts will be reinforced in other activities later on in the course.

Cooperative learning in a computer classroom, helps students to develop teamwork skills. There are many different ways to apply cooperative learning [e.g., Keeler & Steinhorst (1995), Dietz (1993), Garfield (1993)]. In the PACE approach, the type of teamwork that is implemented depends on class size and classroom setup. A typical class size is about 40 students and a computer classroom usually has 20 to 30 machines. When a hands-on activity is conducted in a computer classroom, two students are assigned as teammates, to work together to analyze data, as well as to write up and present a report. A new teammate pair is formed for each activity.

Exercises are essential to help students retain the concepts and skills they have learned. I usually use two types of exercises. One is the simple paper and pencil type of exercise to reinforce students' understanding of individual concepts. The other is a set of working problems related to students' disciplines or current events of interest, such as the home run chase between Mark McGuire and Sammy Sosa. Students need to use a computer to conduct their analyses, and to summarize their findings.

Across each dimension of the PACE model you find technology. It is used for student activities, projects, cooperative learning, and exercises. The technology used ranges from simple calculators to graphing calculators to computer software to interactive multimedia and web technology. Although I have taught introductory statistics courses in two different environments: in a computer classroom setting for the entire semester and in a regular classroom with scheduled time for the computer lab, I prefer a computer classroom environment, which may be either a dedicated classroom or a mobile laptop computer lab. The computer software packages available for my classes include MINITAB, EXCEL, SPSS, and SAS. I prefer MINITAB because it is easy to learn and has all the needed statistical tools.

A typical class

The first day of the class begins with a blurred elephant picture. Students are asked to guess what is on the picture. I then talk about the following two stories. The first story is as follows:

> Once upon a time, there was a King. One day he invited several blind people to his kingdom for dinner. After dinner, the King asked each blind person to touch the animal and give his or her best guess. Someone touched the trunk said it was like a huge snake. Some touched the leg said it was a huge pole, and so on.

The analogy of this story to statistical reasoning and the idea of taking 'subjective' vs. 'representative' sample to estimate an unknown population is discussed. My second story tries to lay a road map for students to make the connection between major topics. I tell students what they will be learning is analogous to preparing for a rescue mission in the Amazon Forest:

> In order to accomplish the mission, you will need to know each tool well and use each tool properly. Then, you will need to simulate similar missions in different environments and learn to anticipate similar patterns and how to deal with them. You will then be sent to the actual Amazon Forest for the rescue mission. This process is similar to what you are going to learn in this class.

The similarity between the progression of different topics in introductory statistics and the mission is drawn and discussed.

These stories are referenced during the semester whenever I see they fit. I find that one of the most difficult learning tasks for students is the connection between major concepts: from the descriptive summary of a sample to the distribution and characteristics of population to the sampling distribution of a sample statistic to estimation and inferential statistics. I use the Rescue Mission story to remind students where they are and what they will be learning next throughout the semester.

My typical class starts with approximately 10 minutes of review exercise by posting one or two problems on the board for students to solve. Here is an example for reviewing the relationship between variability, central tendency, and the shape of a histogram, after introducing the Empirical Rule:

> The price of gasoline is a concern to all of us. A random sample of prices of regular unleaded gasoline from 50 stations in the state is observed, and some summary statistics are obtained: average price = $1.48, median price is $1.45, the sample standard deviation is $.25. Based on this information, draw a price distribution that is more likely to happen, and explain your reason.

Students are allowed to work in groups. Following the review exercise, an activity for introducing new concepts is given. For example, the activity used for introducing the concept of estimation is 'How many raisins are in a box of ½ ounce?' At first, I explain the main objective of the activity is to estimate the unknown average number of raisins in a box by guessing and by actual counts of a sample. I then review the difference between population and sample using this estimation problem, and talk about the terms of parameter and statistic, and the use of a statistic as the estimator of a parameter, especially for mean, variance and standard deviation.

At this stage, I remind students about the Rescue Mission to the Amazon Forest, and that they are ready to be sent on the mission. I talk about the road map of what they have learned and what they are going to learn next. Then each student is given a box of raisins. Before counting the number of raisins, I ask students to guess the number of raisins and write down their guesses. It takes a few minutes for them to count the actual number of raisins in their own box.Two adjacent students are now paired for the data analysis activity. Before the data recording begins, each team has MINITAB up and running. Each student announces his/her guess as well as the actual count, and each team records the variables to a MINITAB worksheet: guess number, and actual count. (A detailed description of the activity can also be found at the website:

stat.cst.cmich.edu/statact/.)

I then post several questions on the screen asking students to compare the distributions, means, medians, standard deviations, and ranges between the guessed counts and actual counts. Then I ask them to draw some conclusions for discussion following the analysis. Each team pair is supposed to decide what graphical displays and data summaries they should obtain and to draw a short summary of conclusions. The discussion session follows, and each team displays their conclusions. I write them on the blackboard and provide a final summary of their conclusions for the posted questions. The differences between the guessed counts and the actual data counts are usually of great interest. The concepts of bias, variability, and outliers are stressed during these discussions. During this stage, the question of whether the sample is a representative sample and the question about estimating the average number of raisins per box by a lower and upper bound are brought up either by some students or by me.

The next stage is group discussion time to discuss estimating the average number of raisins using an interval. Students are asked to determine an interval, relate the level of confidence they have for the interval, and give their reasons. Some hints surrounding the sampling distribution and standard error of the sample mean along with the Empirical Rule, are given on the blackboard. The students' reasons for their proposed intervals are usually quite surprising and interesting. Here are some of the reasons students have given:

- We just made a guess.
- We want to make sure the interval is 100 percent sure.
- By the Empirical Rule, we add and subtract two s.d. to and from the average number to get a 95 percent interval.
- By the Empirical Rule, we add and subtract two s.e. of the sample mean to and from the average number of raisins to get our 95 percent interval.

In a class of seven or eight teams, I usually find that more than 50 percent of the teams are able to find the right solution and give the correct reasons. As expected, I also find many teams give incorrect answers. This provides another opportunity to stress the difference between the standard deviation of individual counts and the standard error of the average counts. My job is then to summarize the 95 percent confidence interval, explain the concept and the application of the confidence interval, and talk about how to generalize a 95 percent confidence interval to other situations. A few working examples are distributed in the class for exercises. Then students are asked to use MINITAB to compute the confidence intervals for the raisin data using the actual counts and the guess counts.

At the end of a class period, I spend about five minutes summarizing what was learned in the class, what will be introduced in the next class, and also remind students to go to the course web site to find out the reading assignment and additional material. I give a daily reading assignment and exercises either from the textbook or on the course web site.

Assessment

Several assessment activities are conducted throughout the semester. As mentioned before, a diagnostic test is given during the first class period to inform students of their basic quantitative background and to learn about the level of quantitative skills that are needed for the class. An online, self-directed quantitative skill exercise is available on the Blackboard course website for students to review and practice these prerequisite skills. The Blackboard course website provides a summary of each students' performance, which allows me to identify and inform those students who do not have adequate quantitative skills so that they might be referred to some remedial work.

Another assessment used at the beginning of the course is a survey I use to assess students' attitudes, learning styles, expectations, and motivations. The survey is recorded using a code (rather than identifying students by name) and is later matched with a post-course survey. In the pre-course survey, I also ask students to write down their learning goals and how they plan to accomplish these goals. In the post-course survey, students are asked to write a letter giving suggestions to future students about how to succeed in introductory statistics. They are also asked to assess themselves, according to how well they accomplished their initial learning goals. At the end of the course I examine the surveys to determine if there has been a positive change in the students' attitudes and to learn students' opinions about class activities. I also use the survey data as part of a collaborative classroom-based research effort I am involved in with colleagues at another institution. I have conducted these pre/post surveys for several semesters and have found the data very useful and informative.

I have relied on three types of assessment to evaluate student learning outcomes: quizzes and exams, student projects, and online assessment exercises. Four tests (including a comprehensive final) are given outside of class time at the Testing Center at CMU. Questions deal mostly with concepts and the relationship of these concepts based on different contexts. About one-third of the questions also ask students to give the reasons behind their answers.

Two reports are required for student projects. The first report requires student to:

- Describe the data, the variables and the data sources.
- Describe how the data fits (a) an independent two-sample problem, (b) a paired sample problem, and (c) a regression and correlation problem.
- Conduct graphical and descriptive analysis for each type of data.
- Draw some reasonable conclusions based on descriptive and graphical analysis.

In the second report, students are asked to add inferential analysis to the first report, using confidence interval, hypothesis testing, and regression modeling or correlation analysis, whenever it is appropriate.

The third type of learning outcomes assessment is the online reinforcement exercise. Self-directed quizzes are available on the Blackboard course website for unlimited practice. These quizzes have been developed to reinforce basic skills and help the students apply concepts to working problems in different contexts. These quizzes are not counted into the student's final grade. However, the course website provides a detailed summary of each student and the whole class, on each question and as a whole. This allows me to find out which students take these practice quizzes and to what degree they take them. I can also determine which students did not make use of these online practices. Some benefits from using the Blackboard course website that would otherwise not be possible include (1) it allows students to take the online exercises at their own pace and their own convenience, (2) it provides immediate feedback for me to adjust the pace, and (3) it gives students unlimited reinforcement as needed. It is expected that not every student will take advantage of the online exercises. However, those students who took the online exercises have indicated that these exercises helped them greatly, especially helped them to keep the same pace as the class without getting behind.

Developing an innovative statistics course

Over the years that I have taught at CMU, I experimented with different teaching tools, but had never paid attention to their impact, other than viewing the changes on my student course evaluations. I think a significant change occurred when I became interested in studying frameworks of learning and the psychology of learning statistics, such as changing teaching strategy from teaching-centered to learning-centered, changing assessment approach from summative outcomes assessment to both formative process assessment and summative outcomes assessment, and changing attention from being a good teacher to being a good learning facilitator and coach. I began to read the research literature in these areas, and experimented with a variety of teaching methods in my classes.

Each component of the PACE approach was incorporated into my classes gradually through several years. For example, when I began to use projects in my class, I experimented with two different types of projects for several semesters, one with a topic selected by the teacher, and one where the students could design their own project. When I first began to use project, I assigned a group of students to work on a project that involved posing a real world problem requiring them to collect data. Groups of four students were formed, using the heterogeneous grouping method of Hagelgans, et al (1995). Students were grouped based on mathematical background, computer skills and gender. Some of the project topics we used included:

- Are there price differences among the grocery stores in the city?
- Location, location, location: Is it true that house prices really are location-dependent in the city?
- The crime rate is declining nation wide. Is it true for our town?

Soon I found that these projects were too large for students in my introductory courses, and that it was very difficult for group members to manage the time needed to conduct the study. One surprising reaction was that most students were uninterested in working on these projects, which I thought were practical and exciting. As a result, their work was often careless and sloppy.

I began to try the second type of project which allowed students to select their own, smaller project. I noticed students were more involved with these self-selected projects and were spending time to find projects that were interesting to them. As a consequence, the analysis and reports were better. Many projects were sports-related. Here are some of the projects students selected:

- Comparison of batting averages between Detroit Tigers and Cleveland Indians
- A model for predicting yardage a running back will attain
- A comparison of salaries between administrators and faculty
- The relationship between salary and number of years served at CMU
- Do people drink coffee or not on a hot afternoon?
- A comparison of the level of lead in the blood among children from battery factory workers and elsewhere
- The relationship between IQ test scores and brain sizes
- The age comparison of divorced couples in the USA
- Snowmobile deaths by counties in Michigan.

I find impressive the diverse selection of projects developed by students. By giving them the opportunity to choose their own topic, students seem to be more motivated to find interesting projects and are motivated to analyze the data and determine the results. There are always some students who need extra help or produce sloppy reports, though. On the whole, I find that the enthusiasm and involvement among students are higher by using self-selected projects.

Another PACE component is hands-on activities. I have obtained a variety of good activities either from the web or from various publications or presentations. Most of the activities that I use are neither new, nor my own creation. For example, the activity 'How far are you from home?' and 'How many raisins in a box of ½ oz?', 'The pulse rate change after one minute of exercise' and so on are commonly used activities by many statistics educators. What may be considered unique in the PACE approach is the integration of activities with a paired team approach for analyzing data and the group discussion for connecting the learned concepts with the new concepts discovered in these activities.

As mentioned earlier, computers play an important role in my courses. However, computing technology has been a problem from time to time. Before the Department of Mathematics had its own lab, I basically relied upon scientific calculators and scheduled computer lab time with the Computer Science Department to demonstrate MINITAB. In recent years, most students own their own computers and the university licenses statistical software such as MINITAB and SPSS. In addition, students are much better at using computer technology. For some semesters, I could schedule a computer classroom for the whole semester, and was able to fully implement the PACE approach. For other semesters, I had to rely on computer demonstration in a regular classroom and provided additional sessions to help students use the software. Course web sites are very easy to develop. My course web site and the Blackboard course web site are important communication channels to inform students about the activities and additional material.

In summary, over the years, I have found that hands-on activities conducted in a computer classroom environment using a cooperative learning strategy appear to motivate and engage students to experience and apply their learned knowledge and connect to the new content to be learned. The variety of assessment activities helps me understand how students learn, where they get lost, and how their attitudes change.

During these years of implementing and modifying the PACE approach, I have experienced a variety of difficulties. These include the availability of computer classrooms, the difficulty of finding an adequate textbook, lack of departmental support, and the amount of time required to prepare for classes. However, once the new approach was fully implemented, I found that teaching became smooth and rewarding, and I felt good that more students appreciated my effort and were willing to work hard in the class.

Future plans

Recently I have been trying to implement an online interactive hands-on activity website to move the locally conducted activities online. Students will enter the data online and it will be stored in a database. The database will then be updated automatically. Based on an instructor's directions, students can select a subset of the data for each online activity. The data can be downloaded in an Excel format for local analysis, and it can also be analyzed online. In fact, I and my colleagues were fortunate to receive an NSF/CCLI grant to develop a real-time online database for hands-on activities after the first draft of this article was completed. A list of real-time online activities is available at stat.cst.cmich.edu/statact/. A few years down the road, I envision the web database will be large enough for data analysis contests or for other creative uses in teaching statistics.

Adequate online assessment is another important item on my agenda. Currently I only use the Blackboard course website for students to do online self-directed quizzes. The development of this website will depend on future technology advances.

I plan to continue my involvement in statistics education research. I use a classroom-based research approach. During recent years, I have had the opportunity to conduct several interview studies focused on how students learn statistics. These interviews reveal more questions than answers. For example:

- How do students learn the concept of variation?
- Where do students get lost in the process of building statistical concepts?
- How effective is computer technology in helping students and in what ways?
- How much impact does attitude have on student performance in introductory statistics?
- How can we motivate students to learn statistics?
- What are the hurdles for learning the Central Limit Theorem and the sampling distribution of sample means?

I am trying, along with colleagues, to gather data to answer some of these questions. Once the data are collected over several semesters and from several universities, they may be very helpful in describing student learning of statistics, and may be used to further improve my courses.

References

Anderson, J., Reder, L. & Simon, H. (2000). Applic-ations and misapplications of cognitive psychology to mathematics education. Retrieved May 10, 2002 from act.psy.cmu.edu/personal/ja/misapplied.html.

Angelo, T. & Cross, K. (1993). *Classroom Assessment Techniques—A Handbook for College Teachers.* (2nd ed.). San Francisco, CA: Jossey-Bass Publishers.

Box, G. (1997). Scientific method: The generation of knowledge and quality. *Quality Progress*, 282, 47–50.

Chance, B. (1997). Experience with authentic assessment techniques in an introductory statistics course. *Journal of Statistical Education*, 5(3). Retrieved May 10, 2002, from www.amstat.org/publications/jse/v5n3/chance.html.

Chen, Z. (2002). Analogical problem solving: A hierarchical analysis of procedural similarity. Journal of Experimental Psychology, 28(1), 81–98.

Cobb, G. & Moore, D. (1997). Mathematics, statistics and teaching. American Mathematical Monthly, 801–823.

Dembo, M. (2000). Motivation and Learning Strategies for College Success. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Dietz, E.J. (1993). A cooperative learning activity on the methods of selecting a sample. American Statistician, 104–107.

Fathom Software (2003). *A Dynamic Statistics Software for Deeper Understanding*. Emeryville, CA: Key Curriculum Press, Key College Publishing.

Fillebrown, S. (1994). Using projects in an elementary statistics course for non-science majors. *Journal of Statistical Educ.,* 2(2). Retrieved May 10, 2002, from www.amstat.org/publications/jse/v2n2/fillebrown.html.

Garfield, J. (1993). Teaching statistics using small-group cooperative learning. *Journal of Statistical Education*, 1(1). Retrieved May 10, 2002, from www.amstat.org/publications/jse/v1n1/garfield.html.

——— (1995). How students learn statistics. *International Statistical Review*, 63(1), 25–34.

Garfield, J. & Gal, I. (1999). Assessment and statistics education: Current challenges and directions. *International Statistical Review*, 67(1), 1–12.

Glasgow, N. (1997). *New Curriculum for New Times—A Guide to Student-Centered Problem-Based Learning.* Thousand Oaks, CA: Corwin Press, Inc.

Gordon, F. & Gordon, S. (Eds.). (1992). *Statistics for the Twenty-First Century.* MAA Notes Number 26. Washington, DC: Mathematical Association of America.

Hagelgans, N., Reynolds, B., Schwingendorf, K., Vidakovic, D., Dubinsky, E., Shahin, M. & Wimbish, Jr. (1995). *A Practical Guide to Cooperative Learning in Collegiate Mathematics*. MAA Notes Number 37. Washington, DC: Mathematical Association of America.

Hartman, H.J. (Ed.) (2001). Metacognition in Learning and Instruction — Theory, Research and Practice. Boston, MA: Kluwer Academic Publishers.

Hawkins, A. (1997). Myth-conceptions! In Garfield and Burrill (Eds.), *Research on the Role of Technology — Teaching and Learning Statistics*. Voorburg, the Netherlands, International Statistical Institute.

Hogg, R.V. (1992). Towards lean and lively courses in statistics. In Gordon, F. & Gordon, S. (Eds.), *Statistics for the Twenty-First Century* (pp. 3–13). MAA Notes Number 26. Washington, DC: Mathe-matical Association of America.

Kalyuga, S., Chandler, P., Tuovinen, J. & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology,* 93(3), 579–588.

Keeler, C. & Steinhorst, R. (1995). Using small group to promote active learning in the introductory statistics course : A report from the field. *Journal of Statistical Education*, 3(2). Retrieved May 10, 2002, from www.amstat.org/publications/jse/v3n2/keeler.html.

Lee, C. (1998). An assessment of the PACE strategy for an introductory statistics course. *Proceedings of the 5th International Conference on Teaching Statistics*, Singapore, 1215–1221.

——— (1999). A comparison of students' beliefs and attitude towards statistics between technology-rich environment and traditional lecture. *Proceedings of the 1999 Math/Science Education Technology (M/SET) Conference*, San Antonio, USA, 133–138.

——— (2000). Computer-assisted approach for teaching statistical concepts. *Journal of Computers in the Schools,* 16 (1), 193–208.

Lee, C., Meletiou, M. & Fouladi (1998). Statistical education theme — Teaching introductory statistics at collegiate level. Paper presented at the Third Annual Conference of Research in Undergraduate Mathematics Education, South Bend, IN.

Lee, C., Meletiou, M., Wachtel, H. & Zeleke, A. (2002). The issue of motivation and expectation in the introductory statistics—obstacles and opportunities. *Proceedings [CDROM] of the 6th International Conference in Teaching Statistics*, Cape Town, South Africa.

Lee, C., Zeleke, A. & Wachtel, H. (2002). Where do students get lost: The concept of variation?. *Proceedings [CDROM] of the 6th International Conference in Teaching Statistics*, Cape Town, South Africa.

McLeod, D. (1994). Research on affect and mathematics learning in the JRME: 1970 to Present. *Journal for Research in Mathematics Education* 25(6), 637–647.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistics Review*, 65, 123–165.

Moore, D. & Cobb, G. (2000). Statistics and Mathematics: Tension and cooperation. *American Mathematical Monthly,* 615–630.

Novak, J.D. (1998). *Learning, Creating and Using Knowledge.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Nelson, T. (1997). Cognition versus metacognition. In Sternberg (Ed.), *The Nature of Cognition* (pp. 625–641). Cambridge, MA: The MIT Press.

Pajares, F. (2001). Toward a positive psychology of academic motivation. *The Journal of Educational Research*, 27–35.

Rittle-Johnson, B., Siegler, R. & Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346–362.

Rossman, A. & Chance, B. (2002). *Workshop Statistics — Discovery with Data*. New York: Springer-Verlag.

Rumsey, D. (2001). Teaching introductory statistics in an immersive learning environment—Planet X. Paper presented at the Colloquium, Department of Mathematics, Central Michigan University, Mt. Pleasant, MI.

Scheaffer, R. L., Gnanadesikan, M., Watkins, A. & Witmer, J. (1996). *Activity-Based Statistics: Instructors Resources*. New York: Springer-Verlag.

Schilling, K. M. & Schilling, K. L. (1999). Increasing expectations for student effort. *About Campus*, May–June: 4–10.

Schroeder, C. (1993). New students — New learning styles. In DeZure, D. (Ed.), *Learning from Change* (pp. 83–84). Sterling, VA: Stylus Publishing, LLC.

Shaughnessy, J.M. (1992). Research in probability and statistics: reflections and directions. In D.A. Grouws (Ed.), *Handbook on Research on Mathematics Teaching and Learning* (pp. 465–494). New York: Macmillan.

Sims, R. & Sims, S. (Eds.). (1995). *The Importance of Learning Styles*. Westport, CT: Greenwood Press.

Stage, F., Muller, P., Kinzie, J. & Simmons, A. (1998). *Creating Learning Centered Classrooms.* Washington, DC: George Washington University.

Steffe, L. & Gale, E. (Eds.). (1995). *Constructivism in Education.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Sternberg, R. (1997).*Thinking styles*. New York: Cambridge University Press.

Sternberg, R. & Ben-Zeev, T. (Eds.). (1996). *The Nature of Mathematical Thinking*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Stipek, D. (1998). *Motivation to Learn: From Theory to Practice* (3rd ed.). Boston, MA: Allyn and Bacon.

Utts, J. & Hackard, R. (2002). *Mind on Statistics*. Belmont, CA: Duxbury Thomson Learning.

Velleman, P. (2003). *ActivStats for Excel.* Boston, MA: Addison Wesley.

Wild, C.J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–265.

4

Statistics as a Core Course in Liberal Arts

Jerry Moreno
John Carroll University

It is my responsibility as a teacher of introductory statistics to provide a meaningful classroom experience for my students to become life-long statistically literate citizens.

How I became a teacher of statistics

I graduated from Lehigh University in 1963 with a degree in mathematics. Having no idea what I was going to do with it, I did the natural thing and went to graduate school (Michigan State University). One of my first quarter courses was probability and statistics. I loved it and went to Jim Stapleton, who was the chair of the statistics department, to inquire about possible assistantships. There were none, but he said that if I did well first quarter they would try to find something for me. I did fine, and they gave me an assistantship during the winter quarter. I finished the master's degree and thought "What now?" I wanted to teach. High school was out because I had not taken any education courses. Besides, I was a statistics major now, not a math major.

As I was in a real quandary about what to do, the department asked me to continue my studies toward a PhD. Their program however was very theoretical and very narrow. I could not see myself pursuing a doctorate at MSU and, besides, had serious doubts that I was capable of doing the terminal degree at all. But a faculty member there, Joe Gani, knew of my situation, and for some reason asked me if I would be interested in going with him to the University of Sheffield in the United Kingdom. Professor Gani had been asked to start a graduate program there in statistics. I was married with no kids; MaryAnn agreed to the move. So off we went to Sheffield. Professor Gani could not promise that he would be able to provide me with a lot of new courses to take the first year as he was just beginning the program, but that I could take what was available and work on filling in some details on the solution to the general epidemic model that he had just completed.

In my first year at Sheffield, one of Professor Gani's instructors at the undergraduate senior level left midway through the year. I was asked to finish teaching the course. At the end of the year, I knew for sure that teaching statistics and probability at the university level was what I wanted to do as a career. As my second year at Sheffield was drawing to a close, Chris Heyde (whom I knew at Michigan State and who went to Sheffield with Joe Gani) asked if I would be interested in doing a dissertation under him, but that he was returning to Australia. The offer was very tempting but I decided to return to the states.

At the time (in the 1960s), having the terminal degree was not required at all universities. I was offered a one-year contract at John Carroll University in 1968 to fill in for a faculty member who ended up not returning from his year's leave. I was asked to stay on and am now well into my third wonderful decade here at Carroll. Having been blessed with a few honors along the way, I am very thankful to Jim Stapleton, Joe Gani, and John Carroll University in playing significant roles in my becoming a teacher of statistics.

My course

John Carroll University (JCU) is located in a residential suburb ten miles east of Cleveland, Ohio. We are a privately controlled, coeducational, Jesuit university that provides liberal arts programs in the arts, sciences, and business for 3300 students at the undergraduate level, and 1000 students in selected areas at the master's level including mathematics. We offer only a minor in probability and statistics. JCU places primary emphasis on instructional excellence enjoying an overall student/faculty ratio of 15/1, although elementary statistics classes usually enroll 24–28 students.

Around 250 students per year take *Elementary Statistics I*, the course I address in this writing. Typically they take the course as a way to satisfy the core curriculum requirement of one mathematics course. Psychology and sociology students are required to take Stats I (in fact, psychology majors must take *Elementary Statistics II* as well). Business students have their own calculus-based course taught in the Economics Department. Biology majors take a semester of calculus and a semester of *Statistics for the Biological Sciences* taught by us in the mathematics department.

Elementary Statistics I has been approved as a core course. The core curriculum in the liberal arts of John Carroll University constitutes 45 percent of the undergraduate graduation requirements. In order for courses to be core courses, they must satisfy a number of principles, the primary ones of which are:

- Core courses are designed to open the mind, stimulate the imagination, and develop general mental skills rather than teach vocational skills to prepare students for specific careers.
- Core courses should help students become aware of their own values and to develop a reflective view of life.
- Core courses should stress critical thinking, problem solving, and oral/written expression.
- Core courses should encourage active learning.
- Core courses should introduce students to the common body of knowledge within a discipline.
- Core courses should introduce students to how an individual discipline uses different methods to generate knowledge.
- Core courses should emphasize connections to other disciplines and contain sufficient interdisciplinary aspects to build bridges to other disciplines.
- Core courses are designed to be foundational in nature, introducing students (both potential majors and non-majors) to material of fundamental and lasting significance.

The fact that *Stats I* is a core course influences to some extent what is taught in the course, and especially how it is taught. In addition to my students evaluating the course at the end of the semester for my personal purposes, they may also be asked to do so by the Core Committee. The latter's questions reflect the eight principles listed above.

My goals in *Stats I* are to prepare all students with the ability to appreciate statistical information they read or see in their everyday lives—now and in their future (Moreno, 1998), to prepare those who will use statistics in their research with a basic working knowledge of estimation and hypothesis testing, and to satisfy the core principles listed previously. It has been a bit of a challenge to teach a liberal arts course according to the Core principles, while providing students the opportunities to learn and use elementary statistics.

Overall, I want students to leave my course with a conceptual understanding of what statistics does and why it is needed, and how statistics does what it does. Fundamental ideas that they should appreciate as citizens are: the reason for statistics; the understanding of graphs; the purpose of basic summary statistics; the design of an experiment, observational study, survey, poll; that statistics does not prove anything; uncertainty and the concept of probability (including risk analysis); misleading conclusions implied from newspaper articles headlines; and "How big is big?" "How small is small?"—*the* questions of statistical hypothesis testing (referring to the significance of values of test statistics and *p*-values).

Regarding textbooks, the department requires that all sections use the same text. There are two statisticians in the department so the course is taught by many of my mathematics colleagues. We have used many books over the years, perhaps somehow trying to find the perfect one that balances statistical concepts and statistical methods. Our most recent selections have been Moore and McCabe's *Introduction to the Practice of Statistics* (Moore and McCabe, 1989), Moore's *Basic Practice of Statistics* (Moore, 1995), *Interactive Statistics* by Aliaga and Gunderson (Aliaga and Gunderson, 1999), Utts and Heckard's *Mind on Statistics* (Utts and Heckard, 2004). However, the primary key to the learning of statistics is not a textbook but rather the innovations of the classroom experience. Textbooks provide resource support for what exciting things happen in those 2100 precious semester minutes between instructor and students.

A typical class

The stats lab in our new science center has eight rectangular tables with 32 flat screen computers, one per student. I have touch screen control of the system with immediate access to the internet, VCR, DVD, and a great document camera. The lab has three walls of whiteboard including a "panel" and other tech bells and whistles. There is a table near the middle of the room for classroom demonstration of experiments.

I have 28 class sessions, 75 minutes in length. Two of the classes are for exams. (Although I have tried to implement other forms of assessment, I just do not seem to be able to get away from examinations.) The length of class is ideal since it allows time for presentation (interactive lecture, video, speaker), discovery (activities in class or on the computer, data collection), discussion, and analysis.

I strongly believe that the classroom experience must provide a real learning environment. If students attend class only to take notes without thinking about what they are writing, to collect information, but nothing new has been learned between the time they entered class and when they left, then I have failed them. For sure there will be topics, perhaps most topics, that are not or cannot be mastered in a single class period, but each student in each class session must learn something. On a daily basis, I try to make the classroom experience rich in addressing the principles of a core course, engaging in the presentation of the day's material, and enjoyable.

Generally, every class has the following format. I begin by allowing students to ask questions on assigned homework. Next, I introduce a new topic. Finally, I assess the day's class. I will describe these three components in more detail in what follows.

Questions on homework

I start a class by allowing about five minutes for students to ask questions on the current homework assignment. My classes are Tuesday and Thursday. In order to give students a chance to ask questions in class on homework, whatever was assigned on Tuesday is due Friday at noon, and whatever was assigned on Thursday is due the following Wednesday noon. These assignments are to be placed in a wall box outside my office. Often a homework assignment will contain an outside responsibility such as running a Java applet, or searching the web for a data set, or looking for a media report on some topic. Most of the time I answer students' questions with questions to get them to answer their own queries. This Socratic type of learning drives them nuts, and I get hammered a bit on course evaluations for "he doesn't answer my questions," but in the long run I trust that their learning is far more complete and lasting.

Presentation of the day's topic

I present topics in several ways. Sometime I collect data through a survey or experiment or student activity. Sometimes I introduce a java applet, video clip, or use an interactive lecture.

I prefer that students work in groups of two or three, but I do not force them to do so unless whatever we are doing absolutely requires it. Actually they usually find company to be helpful. It is important that we use real data in investigating a topic whether it is from a printed source or generated in class through survey or experiment or activity. For example, we will do the random rectangles activity (Scheaffer et al, 1996) if the topic is discovering why randomness is important in sampling; the Six Strings activity or the four M&Ms patterns in a Dixie cup for basic probability (www.jcu.edu/math/faculty/moreno/moreno.htm); data collection of their arm spans and heights for discovering relationships; reaction timers (a "ruler" marked in hundredths of a second that records how quickly one catches it between thumb and forefinger when randomly dropped) or perhaps measuring circumference of balloons blown with one breath, two breaths, three breaths for estimation; or measuring how long it takes an Alka Seltzer tablet to dissolve in various temperatures of water to discover nonlinear relationships. The latter is also very useful in discussing scientific protocol as well.

Or, we might play with a Java applet. For example, I do not include much formal probability in the class except for the binomial distribution. I include the binomial distribution because the only natural, intuitive way to introduce the concept of hypothesis testing in my humble opinion is through proportions. Means are too confusing for many students to handle as an introduction to decision making in having to deal with the parent population mean μ_X, the sample mean $\overline{X}$, and the population mean of the sampling distribution of the sample mean $\mu_{\overline{X}}$, while introducing the concept of the Central Limit Theorem! There are few topics more confusing to students than this plethora of means.

Although I do not introduce hypothesis testing with means, we do eventually get through the tangled web. There are a few ways that I have used to sort out the confusion. First of all, something as simple as using a different color on the board (chalk or marker), one for each of the three means, and using the same colors consistently in whatever future classes the topic arises, has helped many students, so they have told me. Whatever works! Another technique is to use Minitab to simulate the situation. That is useful but it is not completely straightforward. I use Java applets in the course, but as I could not find an applet that did what I wanted to see displayed to illustrate the Central Limit Theorem for the sample mean, I had some of our computer science majors develop an applet for me. (See www.jcu.edu/math and click on Links and then Interactive Statistics Package (ISEP). When ISEP opens, click on Lesson 4.)

This applet shows some population distribution with a random sample of size n (shown as n red dots) emanating visually from random positions. The red dots are being collected in a black box that represents the calculation of the sample mean for each sample. Once there is the specified number, n, of (red) random dots in the box, the sample mean (shown as a blue dot) emanates from the sample mean box and falls through space to its appropriate position on the real line. There are different speeds at which this is shown.

In the beginning, students run the applet at the slow speed to be able to follow the process of a random sample being drawn, its sample mean being calculated, and the sample mean contributing to the buildup of the sampling distribution of the sample mean. The population mean, μ_X, is clearly displayed in the parent population distribution, the sample mean box is clearly labeled as calculating the sample mean $\bar{X}$. After the applet has run a few thousand times (students are able to speed up the action once they get the idea of what is going on), the students will be asked to conjecture about the possible value of the population mean and standard deviation of the sampling distribution of the sample mean that they just created. This applet has helped many of my students clarify “the three means confusion.” There are several other applets there that students find useful.

Once or twice a semester, I show part or all of a video from *Against All Odds: Inside Statistics* (www.pbs.org/als/against_odds/). For example, on the very first day, I ask students what they think statistics is, where statistics are used, etc. After I collect their usually quite limited number of examples, I show them the very first video in AAO. Part of their first homework is to describe in detail three situations shown in the video, the difficulties that were discussed, and the exact way that statistics was instrumental in solving the difficulties. There are seventeen situations described in this video, some in more detail than others, and the students are provided with a very good conceptual view of different scenarios in which statistics is used.

After my students have discovered how the use of several graphical techniques unlock the information contained in a set of data, I turn to developing analytical measures or summary statistics that attempt to do the same thing.

Over the years it has seemed to me that my students have not really understood why we were calculating summary statistics for a set of data. Although students could do the calculations, they saw little reason for why they were doing these calculations, other than just to produce a “correct answer.” If students do not really understand what summary statistics are trying to do, then the measures merely become numbers to calculate. To try to make some sense of why summaries are calculated, I ask them to think of a certain person and characterize that person with **one** word, not a phrase, not a sentence, not a paragraph, not a biography, but **one** word.

The person I choose is someone that everyone knows something about, someone with a multi-dimensional personality. A good candidate for a while was former President Bill Clinton. Whoever is chosen must be someone that the students will look at from different points of view so that their characterizations are quite varied. I list the words on the board and try to arrange them into categories or what I call attributes. One class a few years ago suggested the following words for Bill Clinton: “Democrat, powerful, chubby, jogger, womanizer, President, cute, tall, leader, influential, liberal.” We then tried to put these words together according to some general attributes. One example would be,

Democrat	President	cute
liberal	powerful	tall
	leader	chubby
	influential	womanizer

The general respective attributes might be Political, Professional, and Personal. The idea is that each word in a list characterizes the general attribute of the list, but from a different point of view.

I tell the students that when we characterize a set of data, we do so with **one** number; that is analogous to using **one** word to describe a person. The general attributes we are interested in for a set of data are center, spread, and shape. In each attribute, there are different senses in which the single number characterizes the attribute for the data. For example, we develop the single number summary statistics of mode, midrange, median, and mean to characterize the *center* of a data set. But each does its job in a different sense: mode, most often; median, middle position of the ordered data; midrange, midway between smallest and largest; mean, fulcrum or balance between the total of the positive deviations and the total of the negative deviations from it. Also, the mean may be viewed as fair share, the value that each subject would have were each to have the same value.

The idea is that just as we tried to characterize an individual with one word, which did not come close to telling us all that there is to know about that person, even if we stayed in the same general attribute, so it is with a set of data. We see that one number cannot possibly tell us everything there is to know about the set of data even if we stay within the same attribute. Each of these numbers is a summary statistic. We need a collection of summary statistics within each attribute and over different attributes in order for us to learn about the "complete" character of the data— center, spread, shape.

I usually introduce probability in the following way. I show a coin and ask "What is the probability of getting a head if tossed?" "One-half" the students say. "Why" I ask. In practically the same Pavlovian manner in which they respond "$y = mx + b$" (not remembering much else about what the letters mean) to "What is the equation of a line?" when we study regression, they say "because there are two sides."

I then produce a very long nail and ask "What's the probability that this nail will land point up when dropped on a tile surface?" Based on their previous response, they are forced to say "one-half, because there are two possibilities, 'point up' or 'on its side'." Yet they do not believe what they were forced to say. To cement the idea, I have them put four M&M's, two each of two colors, in a Dixie cup. The M&M's form two possible patterns, same colors opposite each other or same colors adjacent to each other. I ask the students: "What is the probability that the colors will be opposite each other?"

Despite the long nail counterexample, some students cling to "one-half because there are only two patterns." So, we experiment. I have students shake their cup well twenty times (emphasizing that "swirling" is not shaking!) and record the number of times each pattern appears. Their data are written on the board and summed across students. With approximately 500 sample results, the fractions are always within a percentage point of 2/3 for adjacent and 1/3 for opposite. I conclude this experiment with a "proof" by choosing four students to come to the front of the room, two with blue shirts, say, and two with white. They stand in relation to one another, acting out the M&M's positions (two opposite each other in a square) as though they were in the cup. To avoid counting rotations, we fix one of the colors (one student) and have the other colors (students) move in all possible arrangements with respect to the fixed one. They see that there are six possibilities, four resulting in the same colors being adjacent and two in the same colors being opposite each other. In this way, we talk about the relative frequency concept of probability as well as the formal structure of a probability model.

Sometimes I introduce formulas when I believe that this will enhance students' understanding of a topic. For example, I take about twenty minutes of class time to ask questions that lead my students to developing the formula for standard deviation "on their own." They already understand the concept of deviation from the mean. (I define the mean as that number which balances the total of the positive deviations with the total of the negative deviations from it. I take them back to their fifth grade textbook that invariably introduced the mean to them as a fulcrum point on their playground seesaw.) Starting with the mean as a measure of center, idea by idea, and often with the use of a hands-on manipulative that I developed, I lead them through what an appropriate summary statistic for the spread of a set of data might be. They tend to come up with something in the ballpark of what we use as the standard deviation, often the mean absolute deviation. So, when we are done tweaking their conjecture to what is "right," they don't view the result as a formula to be memorized, but they understand its components and how the formula is "doing its job." I have them calculate a few standard deviations by hand on homework, but we quickly turn to technology for large data sets.

Class assessment

At the end of class, I give students a couple of minutes to write a "Memo to Moreno." They jot down the topics that made sense to them, the ones that made some sense but they need to work on, and the ones on which they were clueless. This is not a requirement, nor do students have to sign their names, but it does give me immediate feedback on what went well and what did not go well from the students' perspectives. (I believe I first heard of this idea from Fred Mosteller. He called it "Mosteller Minute Papers," or something like that.) I use the information to prepare what I do in the next class. It is amazing to me sometimes how I have felt that some presentation was awesome, but student responses indicate that they did not understand why I was so excited. My enthusiasm for a topic may indeed cloud the ability of students to grasp the main point.

Student assessment

Despite my attempts to make the presentation of material more innovative, meaningful and interesting for students, I have yet to do much beyond giving examinations (individual) and occasionally assigning projects (group) for assessment. I allow calculators and/or the use of Minitab on exams. There is always a mix of items to assess conceptual understanding as well as items that focus on analyzing data. Because technological tools are permitted, the emphasis in the analysis questions is on interpretation of results in the context of the problem and not on ability to compute.

If I want to determine whether or not students know how to calculate a formula, for example, in the case of standard deviation, I ask them to write a brief development of the standard deviation as a measure of spread. I never ask a student to calculate a formula for the purpose of seeing if they are capable of arriving at the correct numerical answer. As I mentioned previously in discussing the concept of measures, I am trying to see if the student understands that a standard deviation is a measure of variation, a single number whose purpose is to characterize how spread out the entire distribution is from its mean. A couple of examples: draw a graph of a distribution of six data points whose mean is seventeen and whose standard deviation is zero; or, compute the standard deviation for the following distribution, say.

```
                 *
     *           *       *
     *   *       *       *       *
 ----a---b---c---d---e---f---g---h----
```

They should be able to tell me how each part of the formula is contributing to the description. For example, "how spread out the entire distribution is from its mean" corresponds to $X - \bar{X}$ for each value of X, and, in order to get a single number, a sum or product of them must be computed if the entire set of values is to be addressed. But we know about arithmetic averages and so the sum, rather than the product, seems to be more appropriate here.

However, we also know that if we take the sum of the deviations, it will be zero by definition of the mean. So, taking absolute values of the deviations or squaring the deviations takes care of that problem. If we take the mean of the former, the measure is called the mean absolute deviation. In talking through the formula, students develop a sense in which this is a measure of spread. Squaring the deviations presents a problem in that the average of the squared deviations is a squared number. If it is to be used to compare the spreads of two sets of data, it is like comparing the areas of two regions, not easy to see which one is bigger. Moreover, if the measure of spread is to be added and subtracted from a measure of center, then the units must be compatible. One solution is easy; take the square root. (The division by $n - 1$ instead of n is a technical problem that my students accept without a lot of fuss.)

My exams will often include a clip or graph from the newspaper to see if my students are able to apply what we have done in class to a real situation. For example, although *USA Today* has a lot of good statistical material in it, its pictographs are often misleading. Students do well in critiquing them and suggesting a correct version. As an aside, I allow students to use two or three sheets of notes for an exam. The truth of the matter they tell me at the end of the course is that the notes seldom really helped them because if they didn't know the material, they weren't able to find the answer on their crib sheets. But preparing the notes forces them to organize their thoughts and course material. The notes also prove to be useful as a refresher for those who have to take *Stats II*. (Similarly, even more extensive notes in *Stats II* help students as they enter their major's methods course.)

I must say that exam results are more disappointing than I am willing to admit. Some of it has to do with the poor quantitative reasoning skills students have. But as the NCTM and state academic standards and principles make their way into classrooms, I am optimistic that students will gain the ability to think quantitatively earlier on.

There are three or four fundamental ideas that students should carry from *Stats I* through their lives. I call these concepts "quad questions" and tell my students that if I see them on the campus quadrangle six months, a year, or later after taking my class I will ask them a quad question, after proper greeting of course. For example, "By the way, why can Gallup get away with sampling only 1500 people to estimate what tens of millions of people are thinking on an issue?" Well, as you may guess, I am joking more than I am serious, but I would expect them to say something about the class activity we had on the subject in which we made a grid of 200 million squares on the white board. They blindfolded themselves as I painted 65% of the squares blue and 35% of them red, randomly selected. I then covered the board with another board that contained 1500 randomly selected cutout squares (sample). They took off their blindfolds and were to estimate the percentage of the under-board (population) that was painted blue. There's more to the story, confidence interval estimation, but hopefully you get the idea. (We don't actually do the activity on the board, but we do something similar to it using transparencies.)

Developing an innovative course

It is interesting to think back on how I have taught statistics over the years. Waiting in the cafeteria line at a freshman orientation a few summers ago, I had a former student tap me on the shoulder, remind me of her name, introduce her daughter to me as an incoming freshman (a far too common occurrence these days!), and said that whenever she sees M&M's, she thinks of me. I asked her what else she remembered about statistics, and she said that it was one of her favorite courses and she did very well in it. I pushed a bit more to hear how useful the course had been for her as a decision tool in her life, how reading the newspaper had become so much more enlightening. Alas, not so. Finally, she admitted that she liked the course because she loved working with numbers and formulas.

The implication was that I taught number crunching, not analysis of *data*—numbers in context. I'm delighted that she enjoyed the course, but clearly she learned very little other than how to crank out correct numbers from computing formulas. Although I began to use a few activities in the 1970s, by far the emphasis was on formula calculations. We taught according to the textbooks, and the texts typically treated statistics as a mathematics course. I'm glad that my former student wasn't bored or confused by it, but surely she was an outlier. My students were very good at calculating the standard deviation, say, but had little clue as to what it was for, or what it meant. I'm sure that statistics had to be boring and meaningless to them, then and now.

During the early 1970s, I recall having received a set of four volumes called *Statistics by Example*, published by Addison-Wesley, (no longer in print), a product of the American Statistical Association/ National Council of Teachers of Mathematics Joint Committee on the Curriculum in Statistics and Probability for Grades K–12, chaired by Fred Mosteller at the time. (Mosteller et al, 1973). I was not sure what to make of the materials because they were so unlike what was in our textbooks. They were filled with real-life problems! The book's preface stated, "We feel that there is a need to explain real-life problems because the student is unlikely to have had experience with any kind of statistics beyond taking averages, and the teacher with any other than the theoretical side of statistics, if that." How right they were!

Unfortunately the volumes were too far ahead of their time, and they never caught on. But the Joint Committee persevered and wrote The Quantitative Literacy Series (QL) of five volumes (www.pearsonlearning.com/dalesey/dalesey_default.cfm click on Mathematics and find 326–350). Although QL was not for us at the college level, the series became a wonderful resource for those of us who wanted to make the introductory statistics course more data oriented and more "hands-on" but didn't really know how to do it. As an aside, QL was actually designed to support the 1989 NCTM *Curriculum and Evaluation Standards for School Mathematics* that encouraged a new approach to teaching statistics in high schools. The approach featured changes in content, teaching techniques, and use of technology.

College texts were a little slow to respond but there was a definite shift in content and presentation with Moore and McCabe's *Introduction to the Practice of Statistics*. Gone were the computing formulas; modern graphs such as the stemplot and boxplot were introduced and used throughout the text; problems contained real contexts and real data; and a full chapter on design of experiments and sampling designs was unique. My mathematics colleagues actually took well to the new approach. It didn't go far enough for me regarding use of activities, but nonetheless it allowed me to include them in a natural way while following the text. My students don't like to stray from a book too much, so whereas the use of activities in early years was a bit of an appendage to the course, it started to become an integral part of the course while using Moore and McCabe.

I have not yet had a former student tap me on the shoulder to say how much they enjoyed the course and how meaningful it has been for them well after the course was completed. But student course evaluations indicate that because they were involved in class activities, they have a better sense of the material. They see the relevance of the material to their major and to studies reported in the media. It is clear that they enjoy the course; time will tell if they become life-long statistically literate citizens.

Future plans

The three major areas I feel an obligation to improve are assessment, making sure the core principles are being achieved, and trying to balance statistical concepts with statistical methods. I know that exams will always be the major part of student evaluation for me. I have included group projects in my intermediate statistics course and my biostatistics course, but not yet in my intro course. I'm not quite sure why not. Part of it is time. But another part may be that in the non-intro courses, I am somewhat more confident that group members are contributing equally, and that they are more serious about what they are doing. Topics are more meaningful. There is more emphasis on experiments than surveys (my bias showing).

In terms of the core course principles, my plans for change result from student responses to a Core evaluation question "To what degree did this course achieve this goal?" on a five-point scale for each of the stated Course learning goals. Student responses have indicated that I have done fairly well on these measures in *Stats I* except for helping them "become aware of their own values and to develop a reflective view of life." Your first reaction might be "how on earth is *Stats I* supposed to do that?" However, upon reflection, I think that if I make an even greater effort to incorporate examples in class that deal with issues that affect my students' lives, and choose a text whose exercises do the same, then student responses on this question may become more positive.

As we are not large enough a university to afford to offer both a statistical literacy course (that focuses on concepts) and a statistical methods course, I have to find a way to blend concepts and methods into one course. I know that it will require that the number of topics covered be substantially reduced. There are an increasing number of sessions at the ASA JSM annual meetings addressing this issue.

Elementary Statistics I is a wonderful course to teach. It can be taught in so many different ways. Before closing, I must mention a very innovative one that I would like to try sometime and that is Laurie Snell's *Chance* course (www.dartmouth.edu/~chance). But no matter how the course is presented, it must be meaningful in order for a student to become a statistically literate citizen; it must be useful for a students' major; it must be interesting and fun as a classroom experience. It gives many students the opportunity to be able to do something they felt they were incapable of doing, being able to think quantitatively. In so doing, it opens their eyes and lives to new and different life-long experiences.

References

Aliaga, M., & Gunderson, B. (1999). *Interactive Statistics*. Upper Saddle River, NJ: Prentice-Hall, Pearson Education.

Moore, D. (1995). *The Basic Practice of Statistics*. New York: W. H. Freeman.

Moore, D. & McCabe, G. (1989). *Introduction to the Practice of Statistics*. New York: W. H. Freeman.

Moreno, J. (1998). Statistical Literacy—Statistics Long After School. *Proceedings of the Fifth International Conference on Teaching Statistics (ICOTS V)*, Singapore, 445–450, International Statistics Institute.

Moreno, J. (2002). Toward a Statistically Literate Citizenry: What Statistics Everyone Should Know. *Proceedings of ICOTS VI, Cape Town, South Africa*, www.swin.edu.au/maths/iase/Post_ICOTS6.html.

Mosteller, F. Kruskal, W., Link, R., Pieters, R., and Rising, G. (Eds.). (1973). *Statistics by Example*. Reading, MA: Addison-Wesley.

Rossman, A. (1996). *Workshop Statistics*. Key College Publishing. New York: Springer Verlag Publishers.

Scheaffer, R., Gnanadesikan, M., Watkins, A., & Witmer, J. (1996). *Activity-Based Statistics, Instructor Resources*. New York: Springer-Verlag Publishers.

Utts, J., & Heckard, R. (2004). *Mind on Statistics*. Belmont, CA: Thomson Brooks/Cole.

5

Teaching a Technology-Enhanced Statistics Course in a Liberal Arts Environment

Robin Lock
Saint Lawrence University

As statistical software has become more powerful and easier to use, we need to spend less time teaching how to use a package while students can engage in more interesting activities to explore concepts.

Becoming a teacher of statistics

I studied mathematics and education as an undergraduate, and was all set to become a high school math teacher, until my first tour of lunch duty as a student teacher convinced me to send off applications for graduate school. While in graduate school at the University of Massachusetts-Amherst, I taught two courses most semesters as a Graduate Teaching Assistant. I found the business statistics course to be a welcome change from the college algebra service courses that were assigned to most Teaching Assistants. Fortunately, the department combined both mathematics and statistics programs, so I could work on a degree based in both areas.

Towards the end of my graduate program I attended a talk given by statistician Mike Sutherland, called "My favorite data sets." This was the first time I really came into contact with data sets and statistical software, and I thought applied statistics looked like fun. This talk spurred an interest in real data sets that has continued to be a focus of my teaching and professional work.

I was hired at St. Lawrence as the lone statistician to oversee the statistics offerings, although several other faculty participate in teaching the introductory course. I occasionally even teach one of the calculus courses. In 1986, I attended SUNY Oneonta's *Second Conference on Teaching Statistics* and had a breakfast meeting with a number of other "isolated statisticians," including Tom Moore of Grinnell College and Rosemary Roberts from Bowdoin College. We discovered a surprising degree of commonality in the issues we faced in developing statistics offerings within a liberal arts setting as part of a mathematics department. This meeting was the impetus for the formation of the Statistics in the Liberal Arts Workshop (SLAW) —a group of about a dozen statisticians from liberal arts colleges that continues to meet on a regular basis to share ideas about statistics education. These conversations have broadened through participation in the American Statistical Association's Section on Statistics Education, the ASA's Isolated Statisticians group, the quadrennial International Conferences on Teaching Statistics (ICOTS) and the Mathematical Association of America's special interest group (SIGMAA) on Statistics Education to provide a rich support network for developing as a statistics teacher. I am very grateful for the insights, advice and ideas of the many individuals who have participated in these groups.

My course

I teach statistics at St. Lawrence University, a private liberal arts college with about 2000 undergraduate students situated in rural upstate New York. I teach in the (newly renamed) Department of Mathematics, Computer Science, and Statistics.

We offer four (and occasionally five) sections of our introductory applied statistics course (Math 113) each semester to a total enrollment of about 240 students per year. This course is the second most commonly taken course on campus (after introductory psychology). The class meets for three hours per week, either with one-hour blocks on MWF or 1.5-hour sessions on TuTh, for a 14-week semester.

Math 113 is the only 100-level statistics course at St. Lawrence. It has no prerequisites and we see students from all sorts of academic areas. Our philosophy is that students (especially at a liberal arts college) should see statistics as an important field in its own right, with applications to a wide variety of disciplines, before studying the particular uses of statistics in a specific discipline. Several other departments (e.g., psychology, sociology, biology, economics) offer discipline-specific quantitative methods courses that build to varying degrees on the material covered in our course.

My course generally meets in what we call a "smartboard" classroom. The room has traditional desk seating for about 35 students, along with an instructor's computer station and projection system. The middle of what looks like three whiteboards at the front of the room is actually the smartboard that serves as the screen for the projector and as a giant touchpad for controlling the computer. A touch on the screen moves the mouse/cursor, a tap produces a mouse-click, and "pens" allow one to write electronically in various colors.

I use PowerPoint slides for basic lecture notes, although additional notes and examples are presented on the side whiteboards during class. The PowerPoint presentations are available to students through a course folder on the "teaching" drive of our campus network and over the web as part of a Blackboard course. Blackboard is a commercial package used campus wide to provide a consistent environment for web access to course materials. The Blackboard materials also include most class handouts, data sets, textbook tables and web links. We use David Moore's *Basic Practice of Statistics* (Moore 2000) for a textbook and both Minitab and Fathom for statistical software. All students are also required to have a calculator capable of doing two-variable statistics calculations. Although we don't specify a particular model, more than half of our students already own a TI-83 calculator. I provide handouts for using TI graphing calculators and help students one-on-one with other models.

A typical class

The initial few weeks of the semester are spent looking at descriptive statistics, first for single variables and then for relationships between two variables. Day #7 (out of 42 class meetings) is about the point that we begin to discuss correlation for the first time. The previous class will have started material on relationships between variables by looking at categorical vs. quantitative relationships through comparative plots and group statistics. That class might end with a first look at bivariate quantitative relationships by constructing a scatterplot of heights vs. weights from student-generated data. These data were collected as part of a general survey on the first day of class using students in all the Math 113 sections. I select a sample of about ten data points from the larger collection, intentionally avoiding any cases that would be identifiable or potentially embarrassing.

As students arrive for class on Day #7 they find the opening screen for the day's PowerPoint presentation on the smartboard. Standard information each day includes the topics to be covered, a reading assignment from the text, homework exercises, and announcements of upcoming activities.

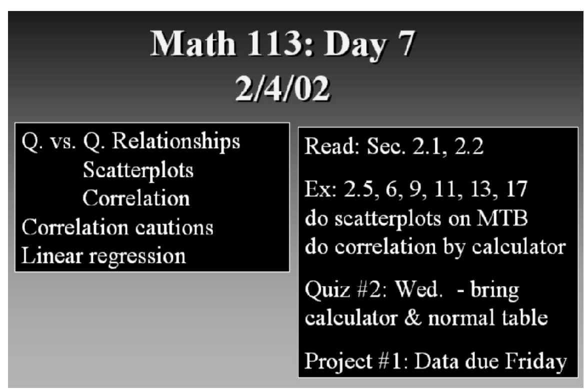


Figure 1. Typical opening PowerPoint slide

After a quick reminder of where we are headed (by displaying a summary slide from the previous class), we review the features that we look for in a scatterplot (e.g., direction and strength of association, linearity, outliers). These ideas are illustrated with a series of scatterplots of data on 1993 model cars (Lock 1993). I pick variables that are relatively familiar to most students (weight, engine size, city mpg, maximum price, median price, and RPM) and ask the students to guess at the type of relationship they might expect as I use the software (either Minitab or Fathom) to produce scatterplots of different pairs of variables. The students view examples of a typical positive association (weight vs. engine size), a negative association (weight vs. city mpg), a curved relationship (engine size vs. city mpg), some outliers (max price vs. city mpg), a very strong linear association (max price vs. median price), and no association (median price vs. RPM). As we display the plots on the screen, I demonstrate how to use the software to highlight and identify unusual points (like the $80,000 Mercedes-Benz). Since it is difficult to take notes on multiple scatterplots, I provide students with a handout (Figure 2) showing a scatterplot matrix (produced by Minitab) that includes each of the plots they have just seen individually.

After students have some experience with assessing associations in a variety of scatterplots, I introduce the concept of correlation as a way to numerically summarize some features of such associations. I provide a formula

$$r = \frac{1}{n-1}\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

for correlation as a product of paired z-scores, although I quickly point out that we will use technology (either a calculator or computer software) to compute correlations in practice. I abandoned teaching the standard computational version of this formula several years ago because those shortcuts offer no insight into what the correlation is actually measuring.

To have the students discover some of the correlation properties, I demonstrate how to use software to compute correlations for each of the original scatterplots from the 1993 car data (basically going down the diagonal of the scatterplot matrix). I suggest that students label the correlations on each of the plots in their copy of the scatterplot matrix (as has been done in Figure 2) as they are computed. I then ask for volunteers to "guess" at the properties of the correlation coefficient. They will readily tell me that correlations should be between ±1, that the sign of the correlation reflects the direction of association, and that stronger associations have correlations closer to ±1, while a scatterplot with no association will have a correlation near zero. I list the properties on the board as they are suggested and then reveal the properties that are pre-set on a PowerPoint slide to see how we did (although I would love to find a way in PowerPoint to reveal the properties one-by-one as they are suggested, as in the Family Feud game show). I have yet to see a class have difficulty in collectively coming up with these interpretations of correlation. I sometimes refer back to the defining formula for further justification that the sign of the correlation makes sense (if X tends to be small when Y is large the z-scores will usually have opposite signs and produce a negative correlation) and

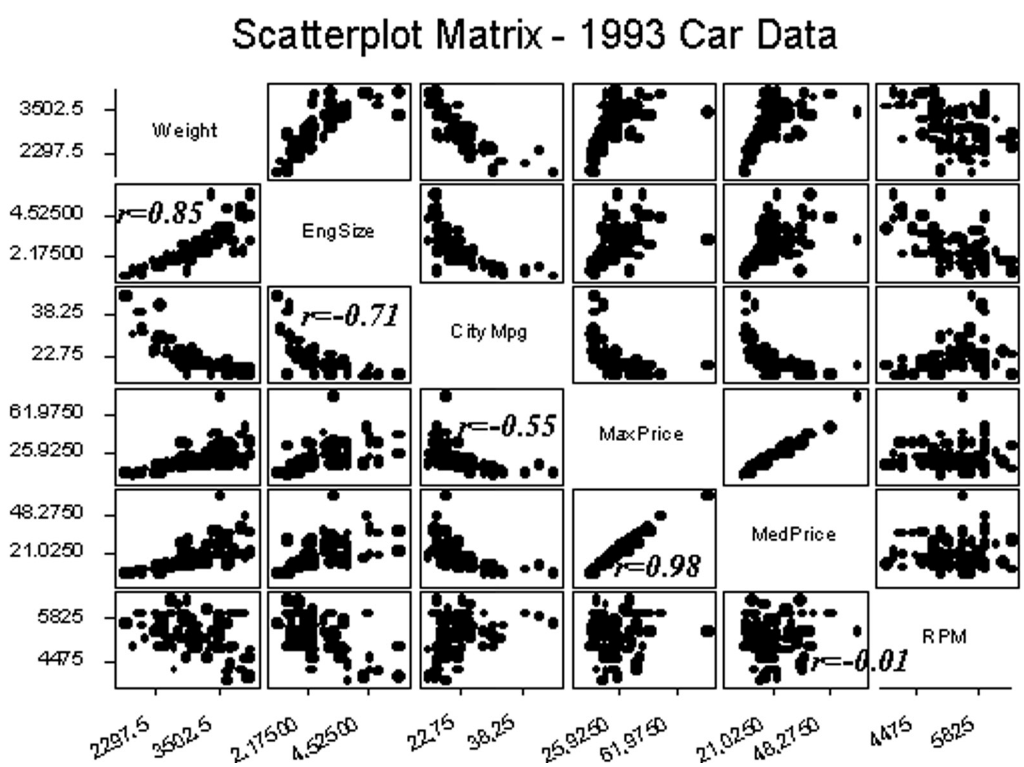


Figure 2

argue (vaguely) that conversion to z-scores helps to constrain the correlation to being within ±1 and independent of the units used to measure the two variables.

To reinforce the connection between the association seen in a scatterplot and the numerical value of the correlation, we look at a Fathom display that shows a scatterplot with about twenty data points and an initial correlation of zero. A slider at the top of the plot allows students to move the correlation to any point between ±1 and the scatterplot changes as the slider is moved to reflect the change in the correlation. I demonstrate this quickly in class and suggest that students can play with it on their own outside of class.

We then go (via a link embedded in a PowerPoint slide) to a web-based Java applet (thanks to John Marden at UIUC) to play the Correlation Guessing Game. Once the applet is loaded we are presented with four scatterplots and four correlation values (as in Figure 3). The task is to match the correlations with the scatterplots. I ask the class for advice on how to make the matches.

They do this with increasing volume as the "game" progresses. The neat hook is that each successful set of four matches contributes to a streak of correct answers. The object is then to build up as long a streak as one can before making a mistake and mismatching one of the correlations. After playing a few rounds of the game on the smartboard with class-supplied input, I show them the leader board that contains the lengths of the longest streaks to date. The default address gives the top 20 scores for the main game, but Prof. Marden has kindly offered to set up school-specific versions upon request via e-mail. Thus our link will go to the SLU version of the game with a leader board consisting solely of students at our school. We start with a fresh leader board in the fall and let the results carry over to the spring semester so that those students can recognize some of their peers on the board.

To encourage students to try the correlation guessing game on their own, I offer a bonus point on the next quiz to anyone who makes the leader board (which includes the top 50 scores in the local SLU version) with a streak of at least 40 in a row. This produces an amazing result as many students will play the game multiple times, first to make the leader board and then to try to move up in the standings. I find e-mails and phone messages when I get to the office in the morning from students who made it onto the board some time late the night before and want to be sure that I note it before someone else comes along and knocks them off the bottom. I even had a former student report that her roommate woke her up at 2 AM because she couldn't decide on a close call (like $r = 0.86$ vs. $r = 0.87$) and needed to complete just two more successful rounds to make it onto the board. I would never have imagined that the prospect of earning a single point (out of 400 for the semester) could motivate students to spend hours matching scatterplots to correlations, but I have probably underestimated the competitive zeal of students who have been weaned on a steady diet of video games.

After we've played enough rounds of the guessing game in class to be sure that all the students are ready to go off and try it on their own, I finish the material on correlation with a discussion of two of the common misconceptions that are made when working with correlations. The first is the tendency to try to infer a cause/effect relationship from a strong association. As examples, I tell them about a strong association between the number of storks living in a village and the number of babies being born there (where the lurking variable was apparently the amount of garbage being sent to the dump as the town grew, providing a more abundant food supply for the storks that dined there) and a strong association between teacher's salaries and alcoholism rates as both increased through the 1960s and 1970s.

The second correlation "caution" is to avoid concluding that a zero correlation means the two variables are unrelated. To illustrate this point I draw a circus cannon on one side of the board and a net to catch the human cannonball on the other. The obvious question to consider is how to predict the distance to the net (Y) from controllable factors such as the angle of the cannon (X). What would a plot of Y vs. X look like if we generated sample data for various angles? Students (with some amusement) can envision the extreme cases of what would happen with an angle of 0° or 90° and will tell me that the maximum distance should occur at about 45° so we can sketch out a plausible quadratic-shaped scatterplot. They readily appreciate that the correlation would be near zero for this plot even though the two variables clearly have a close relationship. This emphasizes the fact that we use correlation to measure the strength of the *linear* association between two variables.

To conclude the day (if time permits), I bring back the original weight vs. height scatterplot that we looked at earlier, ask students to write down a guess as to the correlation, and then use software to compute the actual value. To prep them for the next class (on linear regression) I ask for a volunteer to draw a straight line (on the smartboard) that seems to summarize the relationship between these two variables. I solicit students' reactions

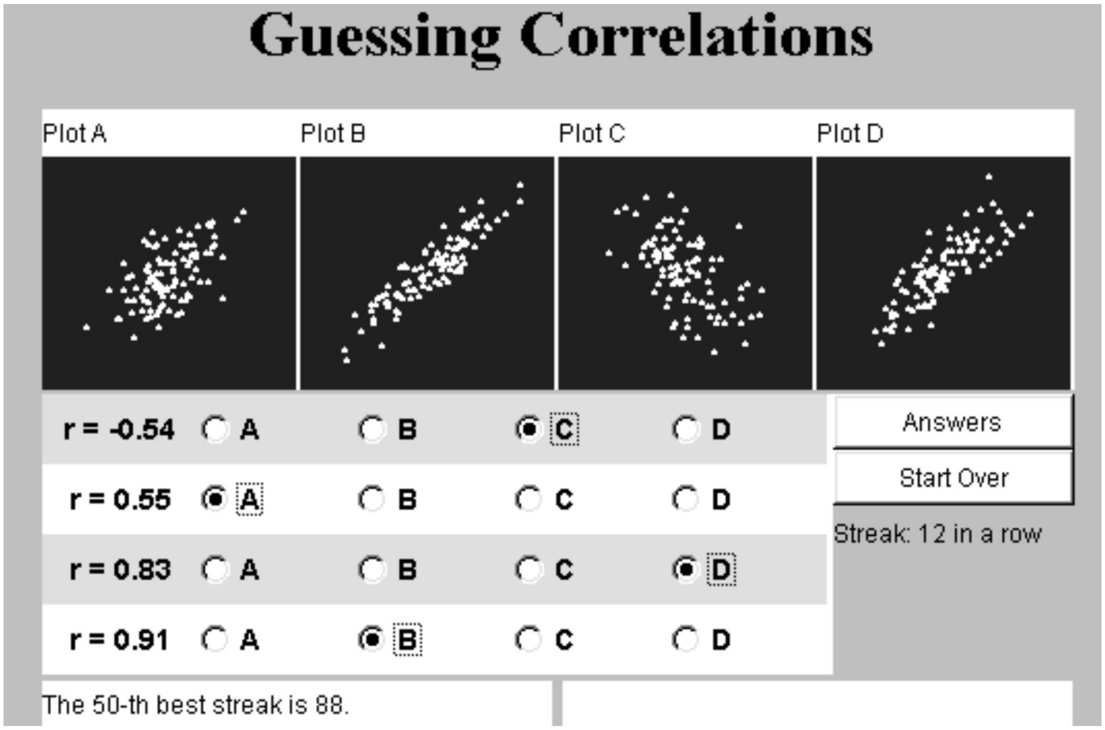


Figure 3. JAVA Applet for matching correlations with scatterplots

to this line, always finding someone who thinks it should be steeper/shallower or shifted up/down, and use that disagreement to tell them that we'll look at criteria for assessing how well a line fits and techniques for finding the "best" fit line at the next class.

Assessment

Although I assign homework problems from the text for each class, I never collect student solutions and rarely discuss the problems in class. I view the homework as a means for students to practice and verify for themselves that they understand the material, and leave it as their responsibility to get help as needed from whatever source (fellow classmates, student solutions manual, evening undergraduate TA's, course instructor, etc.) that works for them. The main means of routine assessment feedback for me is through frequent (roughly weekly) half-hour quizzes in class. I also give two more thorough exams and a comprehensive final exam. The exams given during the semester are scheduled in the evening so that all of the introductory statistics sections take an exam (but not necessarily the same exam) at the same time and we can allow ample time for students to complete exams with no time pressure.

While quizzes typically focus on just two or three concepts and are closed-book, I allow students to bring a limited number of self-prepared note sheets to the evening exams and the final exam. I have found that the process of preparing their "cheat sheets" is a valuable aid in reviewing and organizing course material and minimizes the memorization anxiety. I have even been tempted to grade their sheets rather than their exams, since anecdotal observations would indicate a high correlation between the quality of the note sheets and performance on the exam.

To assess a different group of skills, especially ability to use the computer software and to write coherently about statistical findings, I use three project assignments that are spaced throughout the semester, roughly timed so that they can be graded and returned before each of the exams. These all involve students collecting their own data on topics of their choice and producing some analysis in conjunction with the topics covered at that point in the course. Each student can work on a project individually or may team up with another student to do a joint project. The latter option helps with the grading burden (especially in a semester when I teach 2 sections with 60+ students) and encourages the students to carry on a statistically-based conversation with a peer. Partners turn in a single report and receive a common grade based on essentially the same criteria as individual projects, although I tend to be a bit stricter about careless errors for a project that has been produced by two students.

Generally, about half the students choose to work with a partner and the other half want to work alone on their own topic.

The first of the three projects is the most loosely defined. Students need to find some existing data (producing the data is a topic for later projects) and describe it. They are required to have at least 25 data cases with information on at least two quantitative variables and one categorical variable. An ever-increasing proportion of students find their data on the web with sites such as the *Data and Story Library* (DASL), the *Journal of Statistics Education* (JSE) data archives, or any of the wide variety of sites with sports data being among the most popular options. Students have about a week to find data and get it entered into software (either Fathom or Minitab) and turn in a data sheet with variables and data source identified. I check their data to weed out any that are clearly not appropriate and then they have an additional week to complete the write-up in which they are to summarize and describe the interesting features of their data. Since the format is left intentionally vague, I provide copies of some previous projects on reserve in the library.

Project #2 is a more focused assignment that includes making inference about the mean of a single variable. Students may again use published data, but may also (and most frequently do) generate their own data through an experiment or class survey. Project ideas must be approved in advance of collecting the data and surveys are restricted to the statistics classes— an unfortunate limitation that is necessary due to time constraints on getting approval for wider surveys through the human subjects review process. Once students have collected their data, they summarize their sample (graphically and numerically), construct a confidence interval for the mean, test a hypothesis (using a guess that they made when getting the idea approved), and split the data into two subgroups by some pre-determined factor (e.g., males vs. females) to test for a difference in means.

The final project is based on linear regression where the students choose a variable they would like to "predict" and a set of potential predictor variables. A number of students, who chose particularly interesting data for the first project, find they can return to the same topic and data for Project #3. Part of the project is a fairly extensive analysis of a simple linear regression model based on a single predictor and the rest has students try various combinations of predictors in multiple regression models. Although we do not spend much time on the theory behind multiple regression (and let the computer do all of the calculations), I expect students to roughly interpret individual *t*-tests (to suggest when a predictor might be dropped from a model), assess the overall fit through an ANOVA table and R^2 value, and be able to generate predictions based on a fitted model.

In addition to the three student projects, I use a few "lab" assignments where students all work on a similar topic in small (often randomly assigned) groups outside of class. The first lab generally involves designing an experiment. For example, we might look (in class) at how many drops of water can fit on the face of a coin and then ask the students to design and carry out an experiment to see how a single factor of their choice (e.g., age, heads vs. tails, water temperature, salinity, Canadian vs. American) might be related to the liquid carrying capacity of a coin. This occurs relatively early in the semester, before we've done any formal inference procedures, so the emphasis is on the experimental design with only informal descriptive analysis. The second lab, (also featuring coins), looks at inference for proportions when flipping, spinning and tipping pennies.

Students have an option to do up to two additional projects that, if completed successfully, will allow me to drop their lowest one or two quiz grades (out of 8 quizzes for the semester). The first of these options mirrors Project #2, only it deals with proportions rather than means. The second option (and the most frequently used one) asks students to find an article in a journal, magazine, book or newspaper that uses statistical reasoning in some way and write a critique about how the statistical issues are handled in the article. As part of the write-up, I ask students to include a couple of statistically related questions that they would want to ask the author(s) of the article. In addition to fostering student awareness of the prevalence of statistical issues in the media and the need for making judgments on what is presented, this extra credit assignment helps me to discover areas that students find interesting (e.g., articles dealing with health issues and exercise science have been popular recently).

The various assessment activities show that most students develop fairly good proficiency for choosing appropriate analyses, doing statistical calculations and interpreting the results — at least while they are taking the course. The projects demonstrate a capacity to put all the pieces together, starting with raw data of their own choosing, proceeding through analysis of the data and finishing with a coherent write-up that draws some relevant conclusions. It is not so clear how well these proficiencies are carried along after the course is over. Many of the students from this course go on to take a research methods course in their own discipline. Anecdotal feedback indicates that general principles (e.g., what should you conclude if a hypothesis test shows a very small p-value?) are retained fairly well while specific details (e.g., how many degrees of freedom for a two-sample t-test based on a pooled variance?) are less likely to be recalled in a later course.

Developing an innovative statistics course

Our current Math 113 course has evolved gradually over the 20 years I have been teaching this course. In the early days it had a much heavier emphasis on formal probability and the only statistical software in use on campus was a mainframe package that was available from a limited number of terminals (none of which were near any classroom). A computer "demonstration" consisted of flashing pre-made transparencies of computer output on an overhead projector. The vast majority of the changes that have occurred in the course content and pedagogy since that time have been encouraged by improvements in technological support for teaching and doing statistics.

A good example of how technology can facilitate change is the way we treat correlation in the sample class described earlier. Before we had widely available (and inexpensive) calculators that allow students to enter paired data and easily find the correlation and coefficients of the least squares line, I felt obligated to cover the computing versions of formulas for those quantities. The students had to concentrate on how to maneuver through the various summations needed to apply the formula to a single example and wouldn't have time to discover the broader interpretations that follow from looking at lots of examples where the computations are done automatically.

As statistical software has become more powerful and easier to use, we need to spend less time teaching how to use a package while students can engage in more interesting activities to explore concepts (such as the correlation slider in Fathom). Universal web access has allowed instructors to share applets (such as Guessing Correlations) and datasets (such as the 1993 car data). A student population that comes to us with much more experience in using computers has also facilitated more extensive use of technology. All but a handful of my students have access to a computer and the campus network in their dorm rooms and those that don't have lots of options to use computer labs scattered about campus. Most students need little extra instruction to cut and paste output and graphics from one of the statistics packages into a word processor to produce better-organized and more readable project reports. They are increasingly willing to communicate with me via e-mail and have no hesitation if I suggest that they send me a copy of their data as an attachment.

The downside of software evolution is the need to regularly revise how it is used in teaching. For example, my class is currently in the slightly awkward position of using two statistics packages, Fathom and Minitab. We tend to use Fathom for demonstrations and exploring statistical concepts, while remaining with Minitab for doing analyses for projects (although students have the option of using either and more are choosing Fathom for projects also). Unfortunately, Fathom does not yet have all the capabilities I need for the course (e.g., multiple regression), so, although it may be a superior teaching tool for basic concepts, we will still need to use Minitab for the near future.

St. Lawrence has provided good support for implementing new teaching technologies. Statistics classes can always meet in classrooms with computer projection facilities (although not necessarily a smartboard). Our largest computer classroom with stations for each student has only 25 seats, so the Math 113 classes are often too large to fit, but any of the upper level statistics courses can meet in a room with a computer for each student. We run an integrated network for the whole campus so that the software the students see in the classroom will operate the same way in the public computer labs and any dorm room. With a couple of clicks from my office machine I can make a dataset or class document available to all of my students.

Having access to the PowerPoint slides outside of class has been one of the most frequent positive comments in course evaluations. I was worried at first that making them available would serve as a disincentive to class attendance, but that problem has not materialized. In fact, attendance seems to have improved. Students quickly realize that much of what goes on in class, especially the examples worked out on the sideboards, is not captured by the PowerPoint summaries. Students do seem to appreciate that the main ideas will be available after class, which reduces some of the anxiety about getting everything down in their notes. I was also concerned that having notes prepared in advance would feed one of my bad habits of trying to move too quickly through the material. Interestingly, the effect has been the opposite. Since I'm not as busy writing on the board, I've found I have more time to pay attention to what the students are doing and better able to sense when they are ready to move on. So the frequency of "goes too fast" comments in my student evaluations has actually decreased significantly since using PowerPoint. Finally, students seem greatly relieved at not having to struggle so hard to read my handwriting on the board!

Future plans

As a short-term goal, I would like to explore how the Blackboard system might help enhance instruction, especially through the communication options (virtual classroom, discussion boards, chat rooms, group project areas) that have been underutilized to date. This process should be easier as more instructors on campus establish a Blackboard presence for their courses and students get more accustomed to using the system as a routine part of their coursework. I will also be continuing to develop Fathom applications to target specific statistical concepts, refining the PowerPoint presentations, and monitoring the web for useful material.

Over a somewhat longer (five to ten year) time frame, I would like to see whether the "lab" activities that we do now could be expanded into a more formal lab component for the course. This would require additional meeting times and would pose some significant logistical problems to implement, but would help relax some of the tension that always seems to exist between covering the important topics and allowing students to apply the tools they are learning and discover their own connections.

Beyond the introductory course, our recent hire of a second statistician will allow us to broaden the range of courses that we offer to support our statistics minor and offer our existing upper level statistics courses on a more regular basis. We would especially like to be able to offer a good variety of courses that follow Math 113 (e.g., applied regression, experimental design, nonparametric statistics) and would be of interest to students from other disciplines who are inspired to study more statistics after taking the introductory course or students who might come to campus with Advanced Placement credit in statistics.

References

Data and Story Library (2003). Retrieved October 16, from lib.stat.cmu.edu/DASL/

Guessing Correlations Applet (2003). Retrieved October 16, 2003, from www.stat.uiuc.edu/~stat100/java/GCApplet/GCAppletFrame.html

JSE data archive (2003). Retrieved October 16, 2003, from www.amstat.org/publications/jse/

Lock, Robin H. (1993), 1993 New Car Data, *Journal of Statistics Education,* 1(1). Retrieved October 16, 2003, www.amstat.org/publications/jse/

Moore, David (2000), *Basic Practice of Statistics* (2nd Edition), New York: W.H. Freeman.

6

Teaching a CHANCE Course

William P. Peterson
Middlebury College

The main lesson I have learned is this: To become a teacher of statistics is really to become a learner of statistics.

Becoming a teacher of statistics

Although the trajectory that led to my current position at Middlebury College has not been straightforward, since my undergraduate days I had always envisioned myself teaching at a small liberal arts college. As a mathematics major at Dartmouth College, I became fascinated with probability models. I am indebted to Laurie Snell for introducing me to this field, and also for his wonderful example as a mentor to undergraduates and junior faculty. I began my graduate work at Stanford in mathematics, but soon discovered that a great deal of probability modeling was happening in various allied departments. I earned a Master's degree in Statistics, and finally completed my doctorate in Operations Research, writing my dissertation on queuing networks.

After graduation, I joined a software company as the in-house probability specialist, helping to develop an expert system for manufacturing planning. I have come to call this four-year experience the lab course I never had in graduate school, where I had instead been immersed in the theoretical nuances of limit theorems for stochastic processes. The job provided a crash course in many statistical methods (e.g., estimation of parameters from messy data, model verification and validation). I was often called on to explain what our software did without using statistical jargon or mathematical notation, and that experience has served me well in teaching general education courses at Middlebury.

When my software adventure ended, I rededicated myself to finding a position at an institution that emphasized undergraduate teaching. I was fortunate in 1989 to land a position at Middlebury. In addition to teaching advanced electives in probability, statistics and operations research, one of my early assignments was to teach the introductory statistics course for non-majors. There was no parallel course in my own undergraduate background, so this posed a challenge. Furthermore, the statistical community was in the midst of rethinking the focus of such a course, driven both by the influence of the computer and by pedagogical reform efforts that stressed statistical thinking, real data and active learning (see Moore, 1997). When I began attending the Joint Statistical Meetings in the early 1990s, I discovered a thriving Section on Statistical Education. After attending a few of their sessions, I knew that I had found kindred spirits. Among other things, I learned that they had already thought through the plight of the lone statistician "isolated" in a department of mathematics, and had established an e-mail discussion group to promote communication.

In spring of 1996, I attended the charter meeting of a group called the New England Isolated Statisticians. Our annual gathering has effectively become an extended department meeting, where some twenty academic statisticians have shared teaching strategies, research ideas, and models for liaison with colleagues in industry. I am enormously grateful to all of these colleagues for the many contributions they have made to my development as a statistics teacher.

Middlebury is not far from Dartmouth, so I was fortunate to be able again to visit regularly with Laurie Snell. From our early conversations emerged ideas for a new course called *Chance*. As I will describe later, these efforts eventually led to a large collaborative project involving several other schools, with funding from the National Science Foundation.

My course

Chance is a course about how to read the newspaper. More specifically, it aims to make students more informed consumers of news stories that involve statistical reasoning. The course therefore focuses on reading and responding to real-world applications of probability and statistics as they are reported in the news media.

Any daily newspaper contains a variety of obvious appeals to statistics and probability: weather forecasts include the chance of rain, the lottery reports the odds of winning a prize, the sports page gives betting lines on upcoming contests. Read further and other applications soon emerge. The business pages carry predictions about trends in the stock market, consumer confidence, unemployment, and a host of other economic indicators. Health stories report the latest research findings, along with regularly updated expert recommendations that ultimately can bewilder the lay reader (as I write this, conflicting reports about the risks and benefits of hormone replacement therapy figure prominently in the news). Most public policy debates hinge on some form of data interpretation. Is the Earth getting warmer, and, if so, is human activity the cause? How do today's US students compare with those of a decade ago, or with foreign students today? Would nationally standardized tests improve the situation?

It is easy to go on generating examples. Indeed, it is unlikely that you could look at the front page of your favorite paper and not find a story that included some reference to statistics. However, media reporting about data analysis can be inaccurate or incomplete. Stories that explicitly invoke statistics often garble the meaning of technical terms, or glibly generalize findings beyond the context of the data at hand. Perhaps more challenging are stories whose statistical content is not openly reported, but only implied. The *Chance* course is designed to help students read, interpret, and critically evaluate statistical information from these sources.

I am fortunate to teach in a small liberal arts college environment where there are significant opportunities for curricular experimentation. Middlebury College has two teaching formats that have proved amenable to the *Chance* course: the first-year seminar program and the four-week winter term in January.

All Middlebury students must complete a first-year writing seminar during their first semester. Enrollment is limited to fifteen students, and the instructor serves as their first academic advisor. The College gives specific guidelines for the number of pages of finished writing that students are expected to produce, and asks that particular emphasis be placed on revising work in response to feedback. It is by now well accepted that written expression can enhance the learning process in almost every discipline. In a *Chance* first-year seminar, not only is writing being used to teach statistics—statistics is actually being used to teach writing! This presents some challenges, since topic coverage has to be limited to accommodate writing instruction. But there are also some important rewards, as described later in this chapter.

During winter term at Middlebury students enroll in only one course, which has a minimum of eight contact hours per week. Many students use this term to sample subjects outside their majors. Faculty see the opportunity to present intellectually interesting subject matter free of the constraints that might be present in a regular course sequence, where a set body of knowledge has to be covered. Nontraditional teaching methods are encouraged; indeed, I would find it impossible to sustain a lecture style for the duration of the January experience.

All versions of the *Chance* course share similar goals. First, we hope that students will develop an appreciation for the importance of random phenomena in the world around them, and learn to use statistical reasoning in everyday life. Second, we want our graduates to be critical readers of news stories that involve probability and statistics. Finally, we want students to become comfortable with expressing statistical insights in oral and written English. Terms such as odds, significance, correlation, "law of averages" and "bell curve" have made their way into everyday speech and popular publications, but their original, precise meanings often get distorted along the way. It is important for students to be able to develop their own voices while at the same time maintaining a careful grip on technical points.

In the most adventuresome version of the course, up-to-the-minute news reports dictate each day's topics. Laurie Snell and Peter Doyle of Dartmouth have taught *Chance* in almost exactly this fashion, and find that the resulting sense of immediacy always creates the energy and interest to carry the day. For instructors who prefer a bit more lead-time (I count myself in this group!), we publish a regular electronic newsletter, *Chance News*, which summarizes current news stories and suggests class discussion questions. Archives of the newsletter, along with a variety of other materials to support the course, can be found at the Chance web site www.dartmouth.edu/~chance/.

Although I am not constrained to cover any fixed set of statistical techniques, I have found that several broad themes always seem to run through my courses. The first is that numbers never "speak for themselves." When faced with numerical data, people instinctively feel that they cannot talk back. But of course it matters a great deal where the numbers came from, how they were produced, and what the motivation for producing them might have been. This is really a reflection of David Moore's maxim that data are not just numbers, but numbers in context (Moore, 1992).

A second major theme is to understand the implications of variation, and being able to reason in the face of uncertainty. Qualifying a result with a margin of error statement is not waffling; it is an example of "principled reasoning," to borrow a phrase from Robert Abelson's aptly titled book *Statistics as Principled Argument* (Abelson, 1995).

Third is the often repeated warning that correlation is not causation. Throughout their college careers (and beyond!) students will encounter arguments about causality, as when an economist explains why the stock market bubble burst or a political analyst explains the rise of international terrorism. Although these are not always statistical arguments, statistics provides a wonderful conceptual framework for understanding why causality is so hard to establish. The pitfalls of lurking variables and the virtues of good experimental design are important lessons for all students to internalize.

A fourth theme involves learning how to read a graph, and how to recognize graphical abuses in the popular press. The inspiration here is Edward Tufte's brilliant work on visual displays of information (Tufte, 1983), where he describes in sometimes scathing prose how bad graphics inhibit clear thinking. One of my favorite writing assignments asks students to critique three data graphics that they have selected from current news sources (the only ground rule being that at most one can come from *USA Today*!).

My last theme is lifted from Abelson's Rules, enunciated in the preface in the book (Abelson, 1995) cited earlier. His first rule, "Chance is lumpy," is my favorite. As Abelson explains, occurrences such as a run of ten heads in a sequence of coin tosses are not aberrations; they appear naturally as "undigested lumps in the long stream of data" (p. 21). How surprised are we entitled to be when faced with an "amazing" coincidence? At the heart of any statistical inference, we are really asking this: is the effect I think I am observing real, or can it be plausibly attributed to chance? My course emphasizes the importance of building simple probability models and being explicit about the underlying assumptions, so that these questions can be addressed.

A typical class

The quintessential *Chance* class involves reading and discussing a current news story—in real time. If I anticipate that getting started with the story will require some specialized background knowledge, then I may assign reading in advance, but my preferred introduction involves putting the students into groups, and presenting them with a current story and some related questions. I circulate around the room to gently prod discussion, or clear up general confusion. This requires walking a fine line. I don't want to prematurely steer the discussion, but I don't let the groups wander completely off task. When I sense that some degree of closure has been reached—or total confusion has taken hold!—I reassemble the whole class to pool our insights. One of the standby discussion questions from *Chance News* is "What else would you like to know?" A substantial topic might lead to several related ideas for exploration, in which case I may ask individuals or groups to do further research and report at a later meeting.

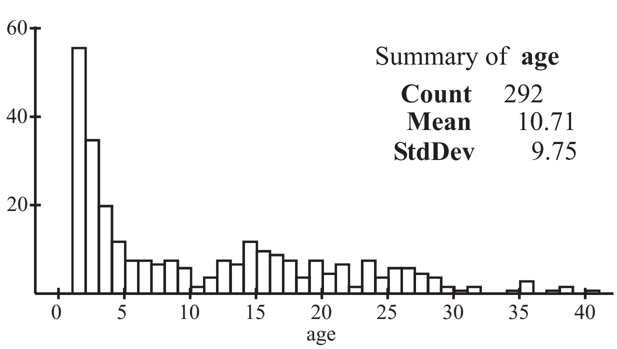
In choosing articles, I look first and foremost for a story with significant current interest. However, I search with an eye towards reinforcing ideas already covered or motivating new concepts. Depending on the topic, I may prepare a hands-on data collection activity, probability game or a computer simulation to illustrate the statistical principles involved.

To illustrate in concrete terms how all of this ideally comes together, I will describe in some detail a recent class experience. Midway through my January 2001 course, we turned attention to a *New York Times* story titled "Opportunity Presents Itself to Kill the Penny" (Tierney, 1999), which had been featured in the July-August 1999 edition of *Chance News*. During the summer of 1999, New York City's bank and stores had suffered an annoying penny shortage. The *Times* wondered how there could be a shortage of something that nobody seems to want, and presented a number of views from people who thought we would be better off without the penny. This is a relatively short article (about 700 words), and I began class by having the students read it together and consider several discussion topics.

The first topic was a study by the Walgreen Company. According to the article, it was found that: "A typical drug store could save more than $2,000 a year by eliminating pennies.… Managers wouldn't have to buy rolls of pennies; clerks wouldn't have to unwrap the rolls and count up the coins at the end of the shift. The average transaction would be 2.5 seconds shorter." What is meant by a "typical" drug store? Do you think this was a controlled experiment? How would you propose to make such a comparison? What is the practical significance of shortening a transaction by 2.5 seconds?

For its part, the US Mint works hard just to maintain the penny supply. The *Times* quotes the Mint's director as saying: "The majority of pennies don't circulate. They make a one-way trip from us to penny jars, sock drawers, piggy banks and the spaces between couch cushions. Two-thirds of the pennies produced in the last 30 years have dropped out of circulation." For the second discussion topic, I asked students how they thought the Mint determined the 2/3 figure.

This provided a natural lead-in to a hands-on estimation activity. In order to have some data available, I had asked the students earlier in the term to begin collecting the pennies they received in circulation. I had them bring these to class on the day I planned to introduce the article. Students went to the board to create a dotplot showing the ages of the 292 pennies that the class had collected. Reproduced here is a histogram of our data.



Actually, before we began plotting, I had asked the students to predict the shape of the distribution, and a consensus had quickly formed correctly anticipating the right skew. Still, their attention was riveted as the plot grew, eventually stretching the length of the board, with the spike at age one spilling over the top. Our classroom is equipped with computer projection, and I took advantage of a later group discussion to prepare the DataDesk plot shown above. However, I would not give up the excitement that was generated by having the group so invested in organizing the data at the board.

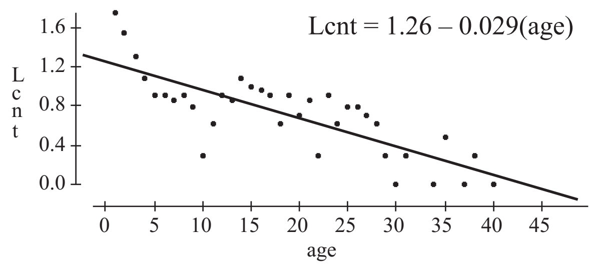
I next asked the groups to consider how to use our data to estimate the fraction of the pennies still in circulation. One group soon indicated that they had a first cut at a solution. Here is a sketch of their argument. Our collection of 292 pennies included 282 from the last 30 years, the time period described by the Mint. Of these, 56 were from the most recent year. Assuming constant annual production, we might have expected to see the same number from each year, for a total of (56)(30) = 1680. Computing 282/1680 = 0.168 gives an estimate that about 17% of the pennies from the last 30 years are still in circulation. This is only half of what the mint reported, so I asked the class what might be wrong with our estimate. After some discussion, they concluded that the count at age one (56 pennies) had too much influence.

I suggested that we needed some kind of model for how pennies drop out of circulation. It seemed natural to the class to assume that a fixed fraction r remain in circulation each year, while $(1 - r)$ drop out. Under this assumption, a recent *Math Horizons* article (Schilling, 2000) formally derives a geometric distribution for the age of a penny encountered "at random" in circulation. This is more sophistication than I expect from my typical *Chance* audience, and I was curious to see where the discussion might head. I was pleased when the class deduced that a fraction r^a pennies of age a years should still be in circulation, so that if M pennies are minted each year, then Mr^a of age a should still be in circulation. Fortunately, we had discussed logarithms earlier in the course, while trying to describe the growth of the College's comprehensive fee. Students were therefore prepared to recognize that

$$\text{Log(Count)} = \log M + (\log r)(\text{age})$$

ought to produce a linear plot (one student recalled having used just such a procedure in a science class). I used the computer to plot Log(Count) vs. age and fit a regression line, as shown below.

For present purposes, the slope value –0.029 is the key figure. Although I had not yet formally introduced the least squares line, most students seemed familiar with



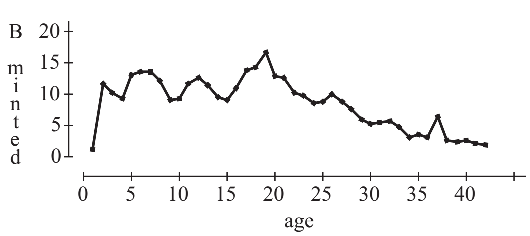
the idea of fitting a line, and were able to produce an eyeball slope estimate of –1.2/45 = –0.03 from the graph. This estimates $\log r$, which in turn gives the estimate $10^{0.029} = 0.935$ for the fraction r of pennies that remain in circulation each year.

The next formula computes the fraction of the $30M$ pennies produced in the last 30 years that are still circulating (with a little encouragement, the class recalled how to sum a finite geometric series).

$$C = \frac{Mr + Mr^2 + \cdots + Mr^{30}}{30M} = \frac{r\left(1-r^{30}\right)}{30(1-r)}$$

Substituting our estimate 0.935 for r gives $C = 0.42$, which is somewhat larger that the one-third figure cited in the *New York Times*.

The students seemed impressed with our efforts. I told them it was now time to look back over the work and ask what assumptions we had made. Discussion was scheduled for our next meeting. The most obvious concern is that penny production might not be constant. Anticipating that this would come up, I prepared the following graph, which shows mintage figures in billions of pennies of each age (the age one figure, corresponding to year 2000, is low because data were available only through January 2000).



Data Source
www.coinfacts.com/small_cents/lincoln_centsmemorial_cents/lincoln_cents_memorial.html (retrieved January 15, 2001).

For ages up to 20 years or so production is oscillating, but further on it seems clearly lower. Students were able to deduce that this would in principle lead us to systematically underestimate the fraction r remaining in circulation. The most conspicuous trend in our own data was a relatively sharp drop-off in early ages. One student wondered if newer pennies might be overrepresented at campus facilities.

I then asked the students to consider what would happen if our sample, supposedly comprising coins from circulation, were in fact contaminated by coins from the penny jars and seat cushions alluded to in the article. (I was motivated here by discussion in "Cents and the Central Limit Theorem" chapter of *Activity-Based Statistics*, which suggests that the mean age ought to be around 8 years. The authors present data with a mean of 10.4, which they consider somewhat large—our class mean was even larger.) Again, it did not take long for everyone to agree that this would over-represent older coins, leading us to overestimate r. Nevertheless, they insisted that they hadn't been on campus long enough to accumulate stashes of old pennies.

By now you may be wondering how the Mint calculates circulation data. The answer appears in a United States General Accounting Office report on the "Future of the Penny" (Gadsby, 1996). We learn there that the circulation rate is calculated "by dividing the number of coins received by Federal Reserve Banks by the number of coins paid out by the Federal Reserve Bank to commercial banks" (p. 7). In 1995, the penny circulation rate was calculated as 34%; in 1991, it was 42%. The report summarizes these findings by saying "almost two-thirds of the pennies produced...are not seen again in circulation." This would appear to be the source of the figure reported in the *New York Times*.

There are several ideas that I hope emerge from the description of this activity. The first is that the *Chance* course draws on many sources for pedagogical ideas, and it is not necessary to reinvent the wheel for every class. I have found that the most successful activities and simulations can be recycled and tailored to current news stories. I had been meaning for some time to try the "Cents and the Central Limit Theorem" activity, and the discussion there was readily adaptable to the circulation story. Naturally, all of this becomes easier after teaching the course several times.

Second, the class did not use a theoretically optimal estimation technique. This is a luxury that I can allow in a course that has no fixed syllabus to cover. We did go through the important stages of proposing and refining a mathematical model, and then retracing our steps to think carefully through its assumptions. To the extent that these lessons are internalized, I would rate the activity as a great success.

Finally, as I have gained experience with the class, I have improved my ability to predict what will be an interesting question for my students. Cholesterol, as I learned the hard way in an early offering of the course, is not something that many students actively think about. On the other hand, the *U.S. News & World Report* college rankings can be counted on to spark a lively debate, and the story is current every fall. I have also discovered that the television drama *The West Wing*, which features story lines inspired by current events, is as popular with students as it is with those of us old enough to worry about cholesterol. Recent episodes have included intelligently written segments on polling, the Census undercount, the *post hoc ergo propter hoc* fallacy, and—you guessed it—eliminating the penny!

Assessment

On the first day of my very first *Chance* course, I ran through the morning newspaper in front of the class, highlighting those stories that depended in some way on statistical reasoning. I then gave an overview of the kind of answers that they should reasonably expect statistics to provide. When we estimate a quantity (e.g., the percentage of voters who favor a certain candidate), we do not merely give a single hard and fast number; instead, we produce a "best guess" along with a measure of our confidence. When we test a hypothesis (e.g., can the SAT exam be coached?), we do not state an absolute conclusion; instead, we assess the strength of the evidence pro or con. In short, I explained, we should not expect there to be a unique "right" answer to any question in this course. Those last words were scarcely past my lips when one student interjected, "Are we allowed to write that on the exam?"

The nontraditional goals of the *Chance* course mandate nontraditional assessment. In the first year seminar, there are three broad areas in which I assess student performance: class participation, writing assignments, and a final "capstone" project.

Regarding attendance, I make it clear to the students that their active participation in discussions is expected; indeed, it is essential for the success of a seminar-style class. Although I do not formally take attendance, I have a variety of strategies to ensure that everyone is participating. When we have a data collection activity, I have each student turn in a worksheet. When I form discussion groups, I have each appoint a secretary to record participants' names and summarize the group's conclusions. When we reassemble as a class, I will often use a randomization procedure to choose a spokesperson for the group (as a pedagogical byproduct, I can demonstrate a variety of randomization procedures). All of these mechanisms taken together allow me to form a composite picture of student participation.

Writing assignments, both formal and informal, are an integral part of the *Chance* course. During a class period, I may ask the students to write a short paragraph responding to an article or describing their reaction to a simulation experiment. While I do not count these towards the writing grade, they provide another component of participation, as well as providing me with ongoing information about conceptual understanding. My thinking about formal writing assignments has evolved over several offerings of the seminar. I have decided that the best approach is to try to model a variety of kinds of writing that they might expect to meet in their college career and beyond.

I have found that a narrative works well for the first assignment, since it does not assume much background. I simply ask students to describe a personal encounter that somehow involves probability and statistics. I stress that it doesn't need to be a success story; in fact, a situation where students wish they had known more might provide a topic for later in the course. This assignment also allows me to foreshadow the idea that a good statistical study ultimately has to tell a story. I am indebted to Abelson for articulating this point.

The next assignment is an expository paper. I often discuss polls and sampling early in a *Chance* course, since I can always count on finding a current poll in the news. The students' assignment is to explain the logic of sampling in terms that their roommate would understand. A college roommate is a good representative of the target audience for *Chance*: a generally educated person with no specialized knowledge of statistics.

At some point in the course, we are bound to encounter a policy debate in the news. This is an ideal time for an argument paper. In my fall 1998 seminar, the press was covering the congressional debate over adjusting the 2000 US Census. *Amstat News* published comments by David Moore, President of the American Statistical Association, who carefully explained the advantages of the statistical approach. These comments were juxtaposed with a scathing column by William Safire, who questioned the motives of the politicians seeking adjustment. We covered various aspects of the adjustment controversy in class, and I asked the students to take a position and defend it in a paper. Later, as part of our rewriting exercises, I asked them to distill their main argument into a 300-word letter to the editor.

I always make time for my favorite assignment on data graphics, which I described earlier. For all of these pieces, assessment is based on demonstrated understanding of statistical concepts through writing. Students learn that understanding a statistical technique well enough to apply it to a news story is quite different from mimicking a numerical example in a textbook. Many are accustomed to short answer exams in mathematics, and assume it is the job of the teacher to figure out what they were trying to say and award credit accordingly. Thus it is important to constantly point out the goal of writing for some broader audience. Different tones and styles are called for as we move from writing for your roommate, to a letter to the editor, and then to a formal research paper.

The capstone experience of the course is a final project on a news story of the students' choosing. The last week of class is devoted to oral reports. I provide foam core poster board for student presentations. I find that this easily rests on chalkboard trays and windowsills around the room on the final day, when the entire collection of posters is exhibited at our *Chance* Fair. I invite my department colleagues to visit and ask questions, and students will often invite friends. The Fair is a nice celebration of the end of the course, and also reinforces the idea that the work is being produced for a wider audience. In the seminar version of the course, a research paper accompanies the final project. I have experimented with a number of intermediate deadlines in the weeks leading up to the Fair, such as asking for a thesis statement, outline, or annotated bibliography.

I am continually refining these assignments in light of experience. I must confess that in that very first *Chance* course mentioned at the start of this section I did give a traditional midterm. This was a mistake, though I did not realize it at the time. The end-of-semester course evaluations revealed that it had created quite a bit of negative feeling. Students noted that the exam was "too focused on details" or that it seemed to distract from the flow of the course. The subtext here, I think, was that the implicit contract of the course had been violated. I had promised in-depth discussions, and the students may have felt reduced to generating sound bites. In any event, the exam has not been repeated. I have steadily gained confidence in my own ability to assess student learning through their papers. A student who has not absorbed the relevant statistical concepts will find it nearly impossible to produce a coherent paper. In *Chance* I am much more interested in seeing whether students can discuss the logic of a statistical investigation than I am in checking whether they can reproduce the mechanics of, say, a *t*-test.

Finally, I should note that I am also interested in how the course affects student attitudes towards statistics. Early on in the *Chance* project, Laurie Snell worried that students arrived to our course willing to believe anything, but left believing nothing! Indeed, as my students gain confidence over the semester, I occasionally find sarcasm creeping into their essays. While it is instructive to pinpoint misapplications of statistics (and there are certainly plenty in the news), it is equally important to prominently feature some exemplary statistical studies during the course. I recently ran across the following very appropriate quote from C.R. Rao: "He who accepts statistics indiscriminately will often be duped unnecessarily. But he who distrusts statistics indiscriminately will often be ignorant unnecessarily" (Rao, 1997).

Developing an innovative statistics course

The Chance course was conceived in conversations that Laurie Snell and I began over a decade ago, in 1990. We were encouraged by the success of the Springer-Verlag magazine *CHANCE* (now jointly published with the American Statistical Association), which had attracted leading statistics practitioners to describe their work in a form accessible to educated non-specialists. We wondered if an introductory course could be developed around stories of current, real-world applications. We received pilot funding from the New England Consortium for Undergraduate Science Education for the summer of 1991, and the project was launched.

The first versions of the course were taught during the following academic year. I offered a first-year seminar at Middlebury in the fall of 1991. Our original conception of the course involved a case study approach, and I structured the seminar around five topics carefully selected over the summer: the 1990 census undercount, the abuse of data graphics in the popular press, the "streak shooting" phenomenon in sports, the statistics of the AIDS epidemic, and the forensic use of DNA fingerprinting. Meanwhile, Laurie Snell and Peter Doyle adopted a more free-spirited approach in the spring of 1992 when they taught *Chance* at Princeton. They were inspired by the remarkable "Geometry and the Imagination Course" that Peter, John Conway and Bill Thurston had created there. Driven entirely by student group discussion, the geometry course had wound its way from the Sherlock Holmes "Which way did the bicycle go?" puzzle to the Gauss-Bonnet Theorem!

Convinced that we were onto something unique and exciting, we successfully applied for funding from the National Science Foundation to develop further materials for the course and to expand the effort to other institutions. Our Dartmouth colleague John Finn wrote for one grant application "*Chance* is an organism." And so it seemed in the early years of the project, as a variety of outlets for the *Chance* concept appeared. Summer meetings at Dartmouth became an annual ritual, where the project team would reconvene to share experiences from the previous year. While we were all having great fun, the universal perception was that the course was very time-consuming to teach. We brainstormed about how to make the load more manageable with the goal of encouraging institutions beyond the project group to adopt our approach. One strategy was to compile profiles of topics that seemed to be recurring in many of our courses, which we posted to our web site. For example, one unit was "Deming and Variability in Manufactured Goods," a problem of enduring relevance, as we now see in discussions of the six-sigma revolution. Another topic was "The Census Undercount." While the news analysis soon became dated, I was able to recycle many of the statistical ideas in 2000. "Using Lotteries in Teaching *Chance*" describes the expected value of a Powerball Lottery ticket, taking into account taxes, the time value of money, and the likelihood of sharing the prize. Whenever the prize pool grows large, these issues are sure to reappear in the news.

An important remaining challenge was how to quickly assemble credible background information on the diverse statistical applications represented in the course. The Chance web site now has video links to twenty-four lectures, featuring outside experts speaking on topics that have been prominent in the news. These include weather forecasting, DNA fingerprinting, epidemiology, the stock market, data graphics, ESP, Benford's Law, and the Bible Codes. The videos have given us access to domain expertise that we could not possibly provide ourselves.

While I find that I rarely show a whole video during a class meeting, I sometimes ask my students to view one for homework as part of a background "reading" assignment. For my unit on Census 2000, I directed students to videos by Tommy Wright of the Census Bureau and David Freedman of Berkeley, who gave excellent talks taking opposite sides of the adjustment debate. I have also found it helpful to show a segment of a video to motivate an issue in class, or just to let the students watch an expert talk through a concept out loud.

One of my favorite videos is "Risk in Everyday Life" by Arnold Barnett of MIT. Professor Barnett is a widely recognized expert on air traffic safety. At one point in the video, he cites the following National Transportation Safety Board statistic, as quoted in the *Wall Street Journal*: "From 1993 through 1996, scheduled US carriers averaged only 0.2 fatal accidents per 100,000 flight hours, less than half the fatal-accident rate for the four year period a decade earlier." He remarks wryly, "There are two things in particular that bother me about this statistic from the NTSB—the numerator and the denominator." Beginning students faced with numerical data often wonder what there is to discuss. A well-timed comment like Barnett's is the perfect antidote to complacency in the face of a quoted statistic.

Throughout this long process of course development, the lifeblood of the project has always been our electronic newsletter, Chance News. I find that I use examples from the newsletter in all of my upper level probability and statistics offerings as well as in Chance. It has even inspired senior thesis projects for several Middlebury mathematics majors. I continue to be surprised and fascinated by the range of statistical questions that turn up every day in the morning paper.

Future plans

We are beginning work on a *Best of Chance News* book, to be roughly organized according to the basic outline of a first statistics course. The volume will present some of the classic stories we have collected over the years, along with classroom activities that have proved especially successful. Conceivably, one could teach a course using exactly these units, but our real goal would be to provide model lessons for teaching in the *Chance* format. Our hope is that instructors would be readily able to adapt our selections to topics of their own choosing.

Recent years have seen increasing calls for required courses in quantitative literacy. An example appeared in the MAA newsletter *Focus* (Steen, 2002) as I was first thinking about this chapter. There is no consensus, however, on the exact meaning of "quantitative literacy." My favorite definition comes from an absorbing essay by George Cobb (Cobb, 1997). Arguing that reasoning is a more appropriate goal than literacy, Cobb identifies four distinct modes of thinking—computational/algorithmic, logical/deductive, visual/dynamic and verbal/interpretive —that need to be integrated in quantitative reasoning.

On the whole, I agree with those writers who have argued that quantitative literacy is not identical with statistical literacy. Nevertheless, as Cobb's essay shows, statisticians have an enormously important role to play in these discussions. The practice of statistics is, by its very nature, interdisciplinary, and the lessons we have learned as statisticians can provide important models for quantitative literacy broadly defined. For my own part, I have applied to Middlebury's college-wide Curriculum Committee to be able to offer *Chance* as a regular semester course in the curriculum. Our distribution requirements for graduation include a course in "Deductive Reasoning and Analytical Processes," a category that sounds tailor-made for the course. I plan to combine the best features of the writing seminar with a more in-depth examination of selected applications. My proposal has been approved, subject to availability of a teaching slot in light of other departmental obligations, and I hope to have more news to report in the not-too-distant future.

References

Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cobb, G. W. (1997). Mere literacy is not enough. In L. A. Steen (Ed.), *Why numbers count: Quantitative literacy for tomorrow's America* (pp. 75–90). New York: The College Board.

Gadsby, J.W. (1996, 16 July). Future of the penny: Options for congressional consideration. United States General Accounting Office, Report number GAO/T-GGD-96-153. Retrieved January 15, 2001 from www.gao.gov/

Gelman, A. and Nolan, D. (1998). Student projects on statistical literacy and the media. *The American Statistician*, 52, 160–166.

Moore, David S. (1992). Teaching statistics as a respectable subject. In Gordon, F. & Gordon, S. (Eds.), *Statistics for the twenty-first century* (pp. 14–25). Washington, DC: Mathematical Association of America.

Moore, David S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review,* 65, 123–165.

Rao, C. R. (1997). *Statistics and truth: Putting chance to work* (2d ed.). Singapore: World Scientific Publishing Company.

Rockmore, D. & Snell, J. L. (1999). To teach, per Chance, to dream. *Notices of the AMS*, 46, 874-877.

Scheaffer, R. L., Gnanadesikan, M., Watkins, A., & Witmer, J. (1996). *Activity-based statistics: Instructor resources.* New York: Springer-Verlag.

Schilling, M. (2000, April). Three estimation challenges: Penny loss, class size, drug use. *Math Horizons*, 22–24.

Sharpe, N. R. (2000). Curriculum in context: Teaching with case studies in statistics courses. In T. L. Moore (Ed.), *Teaching statistics: Resources for undergraduate instructors* (pp. 35–40). Washington, DC: Mathematical Association of America & The American Statistical Association.

Steen, L.A. (2002). Quantitative literacy: Why numeracy matters for colleges and schools. *Focus*, 22(2), 8–9.

Tierney, J. (1999, August 12). Opportunity presents itself: Kill the penny. *The New York Times*, p. B1.

Tufte, Edward R. (1983). *The visual display of quantitative information.* Cheshire, CT: Graphics Press.

Section 2

Statistics Teachers in Two-Year or Community Colleges

7

Teaching Statistics to Under-Prepared College Students

Robert C. delMas
The General College
The University of Minnesota

I believe that one of the key ingredients to becoming an effective teacher of statistics is to have your pedagogy well-grounded in theoretical perspectives.

Becoming a teacher of statistics

My current interest in teaching statistics began during graduate school when I was an instructor for both undergraduate and a masters-level statistics courses. Teaching these courses brought me into direct contact with students' thinking and introduced me to areas of statistics that many students found difficult. I also observed that students at both levels had many of the same difficulties and misunderstandings. I had taken several courses that introduced me to the work of Kahneman and Tversky regarding people's erroneous decisions about chance events (see Kahneman, Slovic, & Tversky, 1982). I was also aware of work in social cognition on the persistence and maintenance of belief systems (e.g., Ross & Anderson, 1982) and of work in science education on factors that affect conceptual change (Posner, Strike, Hewson, & Gertzog, 1982).

Putting together ideas from these different sources, I surmised that many students would have misconceptions about the independence of chance events and that well-formed misconceptions would be resistant to direct instruction. One approach that seemed promising was to create an active learning exercise where students experienced chance events first hand. Another implication of the research literature, however, was that people primarily attend to expected outcomes and either ignore or dismiss unexpected ones. This implied that conceptual change would occur only if students recorded both a prediction for a chance event and the outcome so that both were available for comparison. These hypotheses were well supported by my dissertation results along with the surprising result that students who engaged in the same activity without recording predictions were actually more likely to display misconceptions after the activity (delMas & Bart, 1989).

Upon earning my doctorate, I considered myself an educational psychologist with an emphasis in learning and cognition. Therefore, my first position out of graduate school was as an assistant professor in a psychology department. My instructional assignment included courses in experimental psychology and research methodology where I continued to encounter students' difficulty with statistical concepts, especially reasoning in statistical ways. During this time I kept in contact with Joan Garfield. We shared our observations about students' misunderstanding, hypothesized about the sources of these difficulties, and brainstormed instructional experiences that might be helpful. It was during this time that I developed the prototype that eventually evolved into the *Sampling SIM* software (delMas, 2001). Through our collaboration, I became aware of the numerous areas where cognitive psychology and classroom-based research can be applied to the study of statistics education. Eventually, this interest won out over teaching psychology and led to my current interest and position teaching statistics at the General College of the University of Minnesota.

One of the main influences in my development as a teacher of statistics has been my interest in Psychology (Developmental and Educational Psychology were my respective majors as an undergraduate and a graduate student). I have been intrigued by theories of cognitive development, learning, and cognition for my entire academic career. Many of the activities I use in class stem from the application of theories and research in learning and cognitive development. These theoretical perspectives have led me to develop a more constructivist approach to learning, although my instructional approach is probably better depicted as guided inquiry. Most of my academic research is educational or classroom-based research that investigates the effectiveness of the activities I develop with my teaching/research colleagues. I believe that one of the key ingredients to becoming an effective teacher of statistics is to have your pedagogy well grounded in theoretical perspectives.

Mentoring has also played a significant role in my apprenticeship as a teacher. As pointed out earlier, one of my major influences has been Joan Garfield, who has acted as both mentor and research colleague for over 15 years. I inherited the course I teach from Joan; many of the activities and projects are still based on her ideas. I have also been influenced by the ideas and research of other statistics educators whom I have met through workshops such as STATS (e.g., Robin Lock, Allan Rossman) and Activity-Based Statistics (e.g., Beth Chance, Richard Scheaffer), or conferences such as the Joint Statistical Meetings, the Forum on Statistical Reasoning, Thinking, and Learning, and the International Conference on Teaching Statistics (e.g., Dani Ben-Zvi, Graham Jones, Michael Shaugnessy, Laurie Snell, Pat Thompson, and Jane Watson, just to name a few). As a person develops in their understanding of how to teach statistics, my advice is to be open to the ideas of others and take the opportunity to seek out their counsel and guidance.

My course

I teach an introductory statistics course within the General College at the University of Minnesota. The General College is recognized as a leader in developmental education. According to the National Association for Developmental Education (NADE), developmental education has a theoretical foundation in developmental psychology and learning theory, promotes cognitive and affective growth of all postsecondary learners, and produces instructional programs that are sensitive and responsive to learners' individual differences. In general, students who enter into developmental education programs are under-prepared for college or university work in comparison to students typically admitted into post-secondary programs. Students in the General College are no exception. General College freshmen are similar to other University of Minnesota freshman in that the vast majority are recent high school graduates (about 90%). However, while the average high school percentile rank of freshmen entering the University of Minnesota is around the seventy-fifth percentile, the average high school percentile rank of General College entering freshmen is between the forty-fifth and fiftieth percentiles. Another difference between General College freshmen and the remainder of the University is a higher proportion of students of color. The percentage of students of color who enter the General College is typically around 30%, which is two to two and a half times greater than the University freshman body. A little over half of the students are also first-generation college attendees, so that many of them are unfamiliar with the demands of a college education. The goal of the General College is to provide remedial coursework as needed and to help "develop" students (in one or two years) so that they may transfer to degree-granting colleges at the University of Minnesota.

While the introductory statistics course that I teach is housed within the General College, it also serves students outside of the college. The course is designed to meet the University's liberal education requirement of Mathematical Thinking for students who do not have a strong background in mathematics. Students are typically second semester freshmen or first semester sophomores who do not plan to major in a field with a mathematical emphasis. Students who take the course often see it as a terminal course in their mathematics education.

This statistics course does not lack mathematical content (completion of intermediate algebra is a prerequisite), but more emphasis is placed on developing statistical concepts and procedures for making decisions in applied settings. I teach a four-credit semester-based course that meets twice each week for about two hours each session. The syllabus topics should look familiar to most instructors of introductory statistics: exploratory data analysis techniques (stem-and-leaf plots, histograms, box plots, scatterplots, residual plots), measures of central tendency and variability, probability, sampling and sampling distributions, and inferential statistics (one-sample, two-sample, and paired *t*-tests; contingency table analysis with chi-square).

The overarching goal of the course is to have students develop problem solving and decision making skills through the collection, analysis, and interpretation of data. Students are expected to learn how to analyze and interpret quantitative information, to use statistical thinking, to develop statistical reasoning, and to communicate using the language of statistics. Students are also expected to develop a level of statistical literacy that enables them to critically assess information encountered in the media and other sources.

A textbook by Siegel and Morgan (1988) and a Course Manual that I authored serve as the reading materials. The Course Manual contains almost all of the activities used in the classroom. Many of the activities are paper and pencil exercises that either present real data or have students collect real data in order to perform statistical calculations first hand, carry out procedures, and make decisions. Other activities require students to use the statistical software package, JMP IN, which is a student version of JMP produced by the SAS Corporation. JMP IN is a powerful menu-driven, graphics-based statistical package that covers all of the topics typically presented in a first course. The package is quite interactive and provides "linked" displays. For example, clicking the bar for males in a bar graph of gender highlights all of the respective values in a histogram and a stem-and-leaf plot of footwear (i.e., pairs of footwear owned; see Figure 1). The interface is fairly intuitive making it easy to write clear instructions for each type of analysis.

A typical class

This statistics course relies heavily on active learning through in-class activities. Students meet for a two-hour session twice a week. Every session is held in a computer lab where each student has access to a computer (a typical class size is about 30 students). I spend less than half of a class session lecturing, and usually not more than a half-hour. Lecture time is typically used to introduce a concept, to demonstrate a procedure, or to introduce an activity. During the remainder of a class session, students engage in activities where they learn to apply

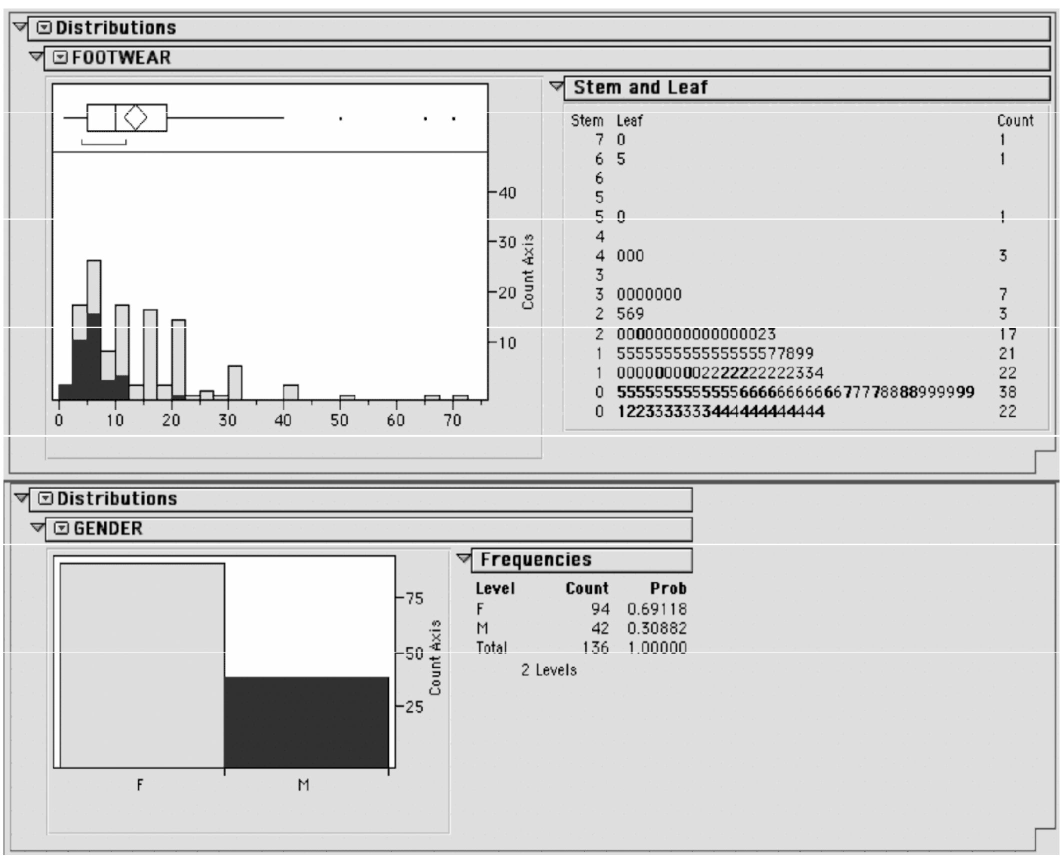


Figure 1. Interactive windows in JMP IN

procedures and definitions or conduct simulation exercises that help them develop conceptual understanding. My teaching assistant and I circulate around the room during an activity to answer students' questions or to inquire how the exercise is proceeding. In this way, we can address most of the questions of each student. Students can move along with material that they comprehend easily and get help when they are genuinely stuck

I will describe four types of activities that represent the different instructional methods used in my course. The first activity involves only simple paper handouts while the second involves student generated data and the use of a statistical software package. The third includes random devices (coins) and web applets, and the fourth describes activities that make use of specially designed simulation software.

My favorite of the low-tech activities is a Sorting Distributions activity that I co-developed with Joan Garfield (Garfield, 2002). This takes place on the second day of the course after students have learned about creating stem-and-leaf plots, which is introduced on the first day of instruction. After reviewing stem-and-leaf plots and showing how they can be used to create histograms, the students break into small groups of two or three students each. One of the group members takes out several pages from the Course Manual on which various histograms are printed. The graphs all represent distributions of real data that cover a wide range of distribution shapes typically covered in an introductory course (e.g., positive and negative skews, bell-shaped, bimodal, uniform, and some that are both skewed and bimodal). The task for each group is to sort the histograms into piles of graphs that are alike according to some set of characteristics, identify one graph in each pile that is typical of the group as a whole, and provide a descriptive name for each pile.

This is a great icebreaker that introduces students to each other. I find students to be very engaged in the activity as I circulate around the room, watch them debate the membership of each pile, and listen to their descriptions. Once each group has completed the task, I go around the room asking each group to identify one of their piles, identify the member histograms, provide their category name, and state what is common to all the member graphs. I usually ask if other groups had the same grouping and what they used as a label. Bell-shaped graphs may be labeled "pyramids" or "mountains," a pile of skewed distributions labeled "ski slopes", skewed bimodal distributions called "city skylines," and uniform distributions viewed as "boring" or "steady-state." After the class exhausts their labels for a particular group, I introduce the statistical term and point out how it matches the labels they created. I believe this activity helps students to remember the different shapes by helping them create personal referents for each type. It also demonstrates that they are capable of creating some of the knowledge that is presented in the course, and that their thinking is valued.

Many of the activities are simply designed to help students become familiar with the statistical software (JMP IN) that is used for data analysis in the course. All of these "get to know the software" activities are based on data from a questionnaire that students complete on the first day of the course. When I began to teach the course, I had each section of students create their own first-day survey. Over several years, I noticed that some of the same questions kept coming up, and that some questions produced very interesting results while others did not. In the name of both sanity and efficiency, I created a single online questionnaire in JavaScript that all students complete on the first day of class (see Appendix 1).

Each student's results are written to a single tab-delimited file that I then post on the classroom server for direct access by the second day of class. Having the single questionnaire allows me to create software lab activities in advance and to have students engaged with the statistical software by the second day of class. One of my favorite questions is one about the number of pairs of footwear owned by a student. A graph of this variable usually produces a positively skewed distribution with outliers. If we examine separate graphs for male and female students, striking differences are revealed in both center and variability. These data also lead to a spurious correlation with other measures such as foot size (which is moderated by gender).

Another of my favorite classes is the session I spend on probability. During the class, I have students perform various activities to produce and record outcomes from random processes. The emphasis of the class is more on the testing of hypotheses than on how to calculate probabilities. For example, almost all of the students will agree that a fair coin will come up heads half the time in a long run of tosses. I then ask for the expected probabilities when the event is defined as two tosses in a row. I usually have a few students who state that there are only three outcomes (both heads, both tails, and one of each) and that the three outcomes are equally likely. Others will state that order matters so that there are four equally likely outcomes. I point out that we now have two different hypotheses that both cannot be correct, but we can run an experiment to see which one is most likely true. At that point, I hand out pennies to each student and they each proceed to generate 20 pairs of tosses, tallying outcomes with respect to the four ordered possibilities. I enter the results into the statistical software and project the statistics and the distributions, and then we decide as a class whether or not one of the hypotheses is better represented by the results.

I present several other activities during the same probability class, each time having the students generate alternative hypotheses, collect data, and judge which hypothesis is supported the best. Some of these activities use computer simulations that I have written in JavaScript, Java, or C++. I wrote the "Where's the Goat" JavaScript applet to simulate the infamous "Let's Make a Deal" game show (students access the applet from the course homepage). In this game, a contestant selects one of three doors. A luxury gift (e.g., an expensive sports car) is randomly placed behind one of the doors while the other two doors have less luxurious gifts (I use a goat, hence the applet's name). Before seeing what is behind the chosen door, the game show host opens one of the other doors to reveal a goat. The contestant is then offered the choice to stay with the original choice or switch to the other, unopened door. At this point I ask the students to vote for which choice they believe will produce the luxury gift most often: staying, switching, or either one (i.e., there is a 50% chance of winning with either choice). The majority of students always vote for "either one," followed by "stay." Very few, if any, vote for the "switch."

The applet allows them to run through trials of switching and staying, producing visual feedback of closed doors, goats, and cars. Students produce a variety of amusing exclamations as they win and lose. The applet tallies total and percent wins for each choice and for total games played. To their surprise, all students find that switching produces the most wins. Most students cannot generate a reasonable explanation for why this is the case. My personal observation is that people tend to only consider the game from the contestant's point of view and do not fully appreciate the role (or knowledge) of the game show host. Before offering an explanation, I point out that regardless of why, the results clearly favor one hypothesis over the others. I then pass out an explanation that attempts to describe the process that produces the results.

A series of activities related to statistical inference uses software that I developed called Sampling SIM (delMas, 2001). The software allows students to easily create a variety of population distributions, quickly generate large numbers of samples of any sample size, view each sample along with summary statistics, and view a distribution of the sample means, again with various statistics reported. Tools are provided to assist students with comparisons and tests of the Central Limit Theorem, among other things. Joan Garfield, Beth Chance, and I have co-developed several activities that use *Sampling SIM* to help students learn about sampling, sample behavior, and the behavior of sampling distributions.

A good example of these simulation activities is one that has students create a scrapbook. Students are given a sheet of diskette labels, three columns by five rows, with a different histogram printed on each label (see Appendix 2). The five graphs in each column represent possible distributions of sample means taken from a single population. For a given sample size and a given population (see Appendix 3), the students' task is to predict which histogram in the column is most likely to result. Once a guess is made and recorded, each student uses Sampling SIM to generate a distribution of sample means for the given sample size. We identify the histogram that comes closest to matching the distribution generated by the computer. Students peel off the identified graph and paste it on the scrapbook sheet. Each square is labeled with the sample size, provides a place to record the guess, and space to record statistics about the distribution such as the average and the standard deviation of the sample means. The same three sample sizes are used for three different populations. (Appendix 3 presents an example of a completed scrapbook.) Students also answer a series of questions designed to help them explore how sample size is related to the shape, center, and variability of the sample mean distributions (see Appendix 4). This is an engaging and very visual way to help students construct the Central Limit Theorem and explore its implications.

Student assessment

I rely on more than one form of assessment to capture student learning across the broad course goals. A student's course grade is made up of points accrued from four types of assessment in the course. The first type of assessment is based on in-class activities. The activities either use real data sets from a variety of sources (e.g., news media, opinion poll sites on the Internet, government Internet sites), have students generate data during class (e.g., exploring probability distributions for ill-defined chance events or drawing random samples from a population of freshman grade-point averages), or require them to collect data outside of class (e.g., information on grams of sugar per serving and cost per ounce for cereals selected from different shelf levels in a large grocery store). Students engage in two or three activities during each class session, although I typically collect only one activity write-up for assessment purposes.

I believe that a great deal of assessment (and self-assessment) actually happens during an activity. Students work on the activities individually, although they are encouraged to consult or work with another student. Instead of lecturing, the teaching assistant and I are free to roam about the classroom and address individual questions. I encourage students to check their work with each other and to ask another student for help before they ask an instructor. They are told to contact an instructor only when they are really stuck. They are also required to try and understand the material for themselves. By the time an activity is completed, most students have had their individual questions addressed. In this way, students receive immediate feedback that supports self-corrections of their procedures and thinking. This system also makes grading the activities fairly straightforward. While there are very few statistical errors, I often find the opportunity to correct or comment on the use of a term, and minor procedural infractions do appear that may have been missed during class. The corrected in-class write-up is always made available by the next class session. Each student turns in around 15 in-class activities by the end of the semester. Performance on the collected in-class activities constitutes 15% of the course grade, which also creates an incentive for class attendance and participation.

Students are required to complete several small homework assignments of six to ten textbook problems. The textbook problems are carefully selected to reinforce the concepts learned in class and to offer students an opportunity to demonstrate what they have learned. A set of textbook problems is selected to accompany almost every class session, and the assignments are due by the class session following presentation of the respective topic. This creates another incentive for class attendance, but also emphasizes the cumulative nature of material presented in the course. Students typically have access to three different teaching assistants and three different course instructors outside of class time to answer questions about textbook problems. I also receive and address many questions by email. Once all sections of the course have turned in a homework assignment, solutions are posted at the course homepage for the students' review. The homework sets are graded by the teaching assistants using a global 0 to 4 rating scale that essentially rewards a student for attempting the problems, showing their work or thinking, and producing reasonable responses. Teaching assistants report that it takes a half-hour to an hour to grade an assignment for a section of 35 students. Performance on textbook problems makes up 10% of the course grade.

In-class activities and homework assignments provide the primary preparation for the course exams. I offer three exams during the term and a comprehensive final examination at the end of the term. I do not believe in requiring students to memorize formulas. They are allowed to bring a single 8.5 by 11-inch sheet of paper with information written on both sides to the exams. Each exam covers two or three topics (about three weeks of material), consists of five or six problems each broken into multiple subparts, and takes students anywhere from 45 to 90 minutes to complete. All together, the exams make up 55% of the course grade: 10% for each term exam, and 25% for the final exam.

Exam items may require students to perform some statistical calculations, but usually in order to produce a graphical display (e.g., median, quartiles, thresholds, and adjacent values for an outlier boxplot). Problems typically ask students to make a comparison or a decision. One example is an item that presents a stem-and-leaf plot along with statistics such as the median, quartiles, mean, and standard deviation and asks students to justify their choices for the best measures of center and variability to use for summarizing the distribution. The Course Manual has a set of about 20 review items for each exam with solutions to all items available from the course homepage. I encourage the students to use the exam review items not only to prepare for the exams, but also as worked out examples for homework assignments.

The fourth type of assessment consists of two student projects that require students to apply statistical concepts and procedures to data they collect outside of class. For the midterm project, students identify three quantitative variables that measure something about their daily life activities and collect daily measurements on the variables for five weeks. Students are free to choose just about anything, but are asked to consider variables that they believe might be related to each other. Some of the most common measures are daily hours of sleep, study time, television viewing, and exercise. Athletic students often choose fluid intake per day to see if it is related to minutes of exercise per day. During the sixth week of the course, one class session is set aside for students to enter and analyze their personal data in accordance with a list of questions that define the project report. The questions are designed to have students demonstrate the exploratory data analysis concepts and techniques acquired during the first five weeks of the course.

The final project instructs students to develop a questionnaire and administer the instrument to a sample from a target population of interest. The project requires students to think ahead about hypotheses of interest and to design the questionnaire so that several comparisons can be made between two groups of people (e.g., men and women, upper- and lower-division students, college and non-college adults). Half of one class session is used for students to brainstorm ideas and get feedback from other students and the instructor. Prior to collecting and analyzing the data, students are required to submit a draft of the questionnaire. After receiving feedback on the questionnaire but prior to collecting the data, they submit a draft of the introduction and methods section for review and comment. The revision process is designed to emphasize the need for clearly thought out hypotheses and clearly worded questionnaire items. The process also illustrates that much of the statistical analysis is determined ahead of time by a well thought out research design. The last week of the course is dedicated to entering and analyzing the questionnaire responses and for review of report drafts by the instructor. These projects provide an alternative way for students to demonstrate their understanding of the course material. The midterm counts for 10% of the final grade, and the final project for 15%.

I have found that the vast majority of students perform very well on the midterm and final projects. About 70% and 80% earn a grade of B or higher on the midterm and final projects, respectively. They demonstrate an understanding of how to take a personal belief and shape it into a question that can be addressed by statistics, to collect data that addresses the question, and to select appropriate statistical analyses. Performance on exams indicates that the majority of students have developed at least a moderate level of statistical fluency from the course. These students can distinguish different types of graphic displays and use information in graphic displays to make appropriate decisions. They understand conditions under which statistics are and are not appropriate. They can identify an appropriate inferential procedure to use with different types of data, follow procedures to draw correct inferences, and relate the practical implications of the results. While this is the case for most of the students, it certainly does not describe them all. About 50% of the students earn a grade of B- or higher, with another 30% earning a grade in the C range. This is a relatively high success rate given that they are primarily under-prepared students in their second or third semester of college who are attempting to learn the challenging material presented in an introductory statistics course.

Developing an innovative statistics course

The statistics course that I teach is highly dependent on the growing capabilities of technology, especially computers. I am a self-taught programmer, and I love to come up with visual approaches to complex statistical concepts. Once I have an idea, one of the greatest challenges is simply finding the time needed to develop that idea into a software program. My focus at the moment is in the further development of Java applets that allow students to explore many of the core ideas in statistics: distributions, center, variability, sampling, sampling distributions, confidence intervals, and p-values. It takes time to learn new technologies (and programming languages) and to keep up with the technologies as they develop. I do find, however, that the time required to create an applet for a specific idea pays off because the approach can usually be generalized to several other areas in statistics.

Trying to make a course more innovative inevitably leads to obstacles. One of the main obstacles I have faced is that a guided inquiry/active learning approach is unfamiliar to many students. It takes students a little while to accept the structure of the class. By the second week, most students are fully engaged and start to gain an understanding of what is required. I am always pleasantly surprised at how many students comment on course evaluations that they had a unique experience in my course, perhaps one of their best learning experiences, and that they wish more courses were taught this way. Many students will comment that not only did they learn about statistics, but they also gained a better understanding about how they learn.

When I develop a new activity for my course, I do not expect to get it right the first time. I have learned that every activity will need to go through several iterations (i.e., over several semesters) before most of the bugs are worked out. Because of this, I try to introduce only one or two new activities in a given academic year. I also let students know when I try something for the first time, or if I am modifying something that was tried before. It is their education that I am “playing” around with, so I think it only fair to apprise them of the situation. I also find that most students tolerate a few little glitches in the activity and often offer useful suggestions on how to make changes

I view the use of each activity with students as a type of classroom research. I begin by stating clear goals for the activity in order to identify assessment items that will provide meaningful feedback. This also relates to my belief that it is important to have the activities grounded in a theoretical perspective. An attempt to wed instructional goals with activities, and activities with anticipated outcomes that are measured through assessment, creates the opportunity to engage in statistics education research as well as develop new instructional approaches. I often find myself moving through a cycle where instructional goals inform the design of an activity, the interaction of the activity and the goals suggest different types of assessment, and assessment results suggest useful changes for the activity. Using this classroom research approach to gather feedback helps me to modify, improve, and validate an instructional approach.

Future plans

There are very few aspects of the way I teach and what I teach that cannot benefit from some improvement. While most students seem to benefit from the activities I use, there is always a small group of students who do not seem to understand an idea or concept, even when they appear to be fully engaged in the instruction. There is always the possibility of making a concept accessible to a larger number of people, and that motivates me to experiment with alternative ways for students to interact with the instructional material (typically through technology). One of my future plans is to continue observations of how students learn, discuss my observations with colleagues, and continue to develop and refine activities to test out our ideas.

Recently, I have attempted to make the big ideas of statistics more integrated throughout the entire semester. For example, variability is a very important, unifying concept in statistics that is essential to understanding many other statistical procedures and concepts. I used to teach a class session or two that focused on variability, assumed that students understood enough to carry on with the remainder of the course, and then referred to variability as it came up in related topics. I used to operate more from the assumption that students would see the connection rather than lead them to make the connection. Having found that students rarely make or understand these connections, I am now convinced that students need more opportunities to develop their understanding of variability, and that this understanding can continue to develop throughout the course. I have started to place more emphasis on variability as one of the major concepts in statistics, not just when it is introduced, but in just about every class session. I believe that students need to understand that variability is more than just a term or the result of performing a calculation. Hopefully, this approach will make topics such as sampling variability and inference more accessible to students, a hypothesis that I hope to explore.

References

delMas, R. (2001), *Sampling SIM*. (version 5), Author. Retrieved October 10, 2003, from www.gen.umn.edu/faculty_staff/delmas/stat_tools/.

delMas, R. and Bart, W.M. (1989). The role of an evaluation exercise in the resolution of misconceptions of probability. *Focus on Learning Problems in Mathematics*, 11(3), 39–54.

Garfield, J. (2002). Histogram sorting. *Statistics Teaching and Resource Library* (STAR), Retrieved October 10, 2003, from www.starlibrary.net/activities/garfield2002.htm

Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Posner, G. J., Strike, K. A., Hewson, P. W., and Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227.

Ross, L., and Anderson, C. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In D. Kahneman, P. Slovic and A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 129–152). Cambridge: Cambridge University Press.

Scheaffer, R. L., Gnanadesikan, M., Watkins, A., and Witmer, J. A. (1996), *Activity-Based Statistics: Student Guide*. New York: Springer.

Siegel, A.F. and Morgan, C. J. (1998). *Statistics and Data Analysis: An Introduction*. New York: Wiley.

Appendix 1

Part of the First Day Questionnaire

1. Who is your instructor? ○ Suzanne Loch ○ Brenda Tiefenbruck

2. Which section are you in? ○ 001 ○ 002 ○ 003 ○ 004 ○ 005

3. What is your gender? ○ MALE ○ FEMALE

4. What is your age in years? [ ] years of age

5. How many siblings do you have including half- and step- ? [ ] siblings

6. How many pairs of footwear do you personally own (this includes shoes, boots, sandals, clogs, etc.)? [ ] pairs of footwear

7. How many body piercings do you have on your body? [ ] piercings

8. Rate your intelligence level on a scale of 1 to 10
(where 1 = Brain Dead and 10 = Rocket Scientist)

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

9. Do you live? ○ OFF CAMPUS ○ ON CAMPUS

10. What is your cumulative GPA? [ ]

11. How many miles do you travel from your current home to campus each day, to the nearest mile? [ ] miles

12. How many minutes does it take you to travel to school each day, on the average? [ ] minutes

13. What type of transportation do you use most often to get to school?

○ WALK ○ CAR ○ BUS ○ BIKE ○ OTHER

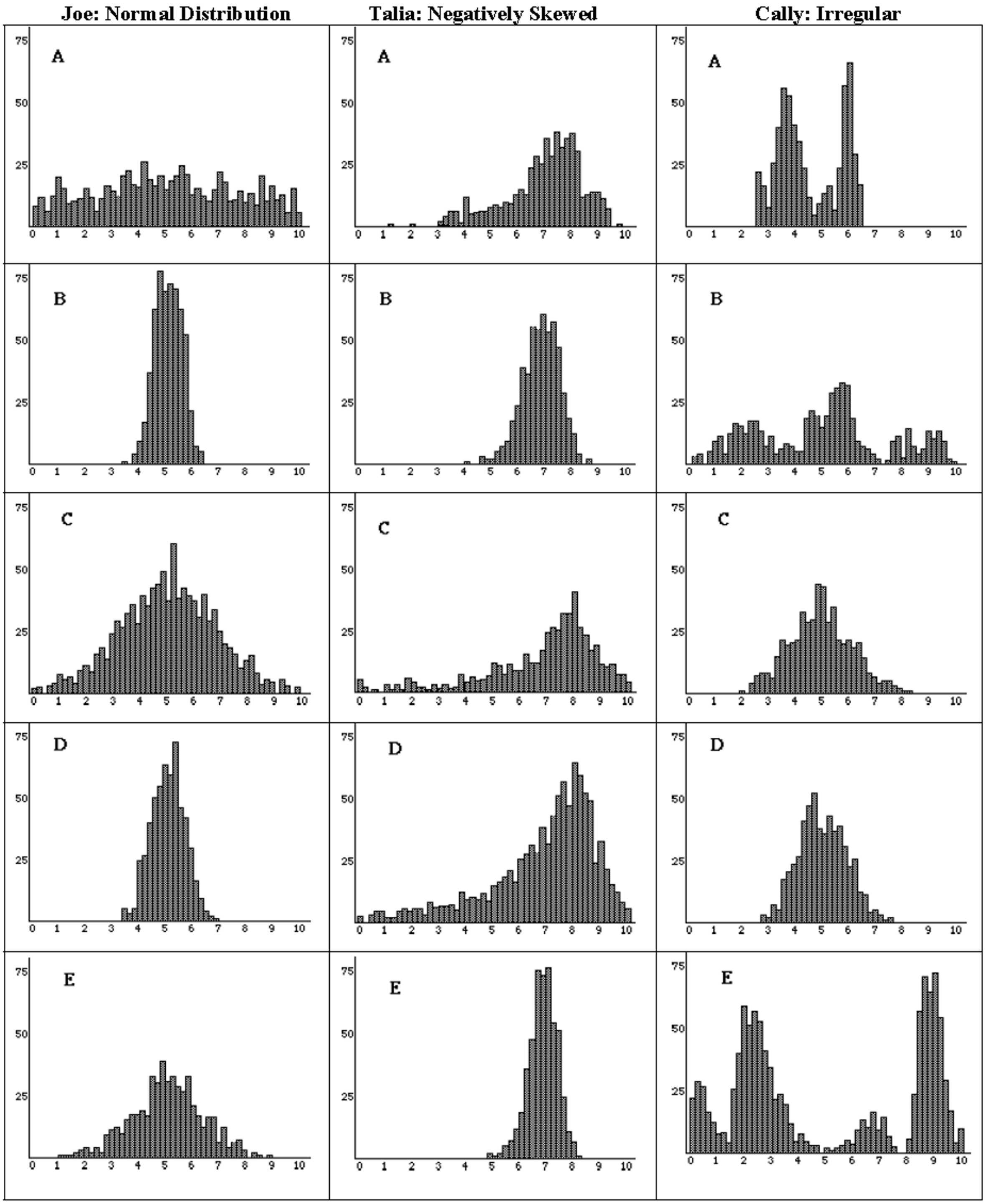
14. How many minutes do you exercise each week, on the average? [ ] minutes

15. How many credits are you taking this semester? [ ] credits

16. How many hours per week do you study, on the average? [ ] hours

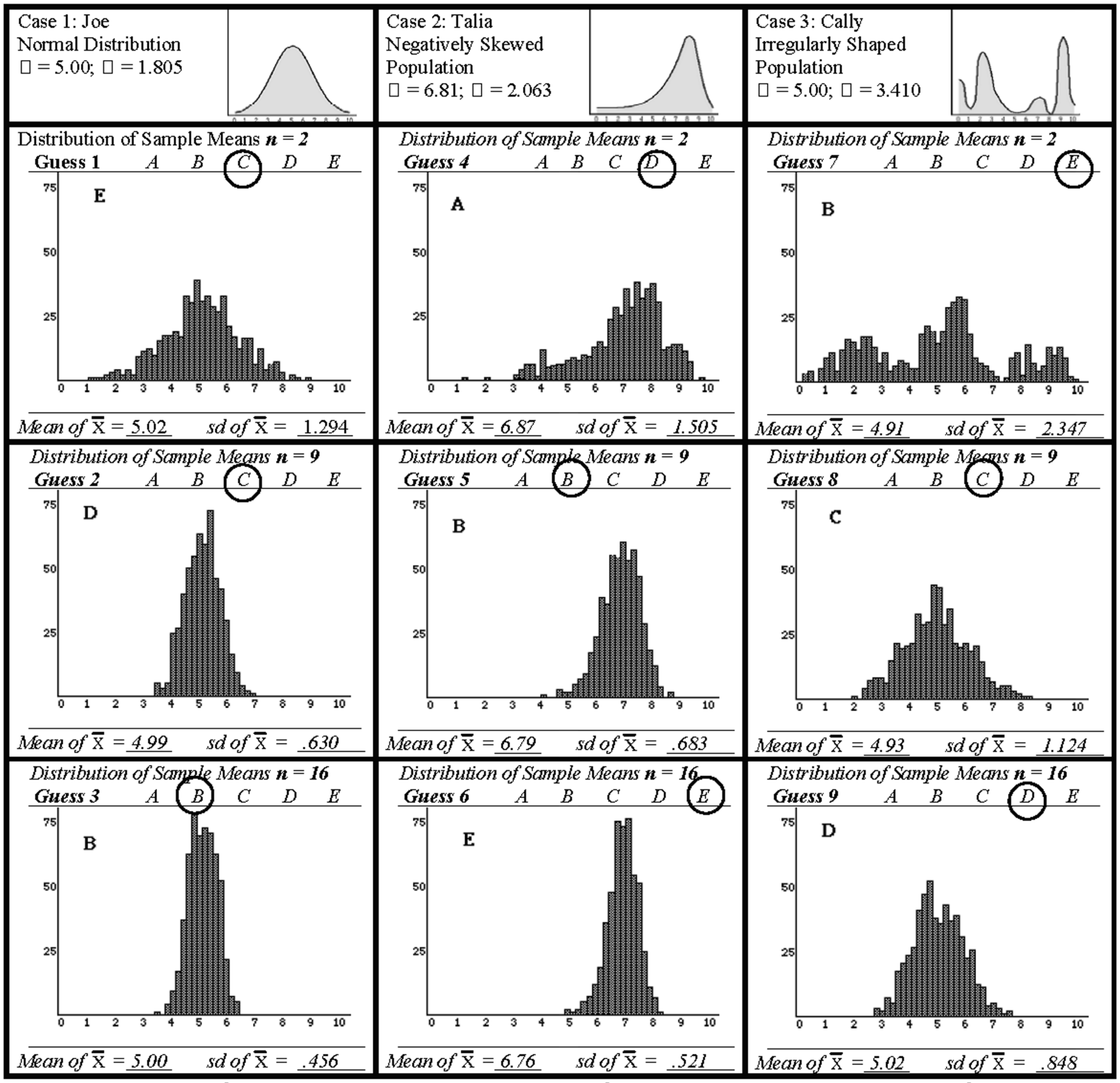
Appendix 2

Sheet of graph labels



Appendix 3

Example of a completed scrapbook



Score: 1 out of 3 Score: 2 out of 3 Score: 2 out of 3

Appendix 4

Questions used to guide inquiry about relationship of sample size to distributions of sample means.

The first set of questions is concerned with the **SHAPE** of the distribution of sample means.	Joe Normal Dist.	Talia Negative Skew	Cally Irregular Shape
For a sample size of $\boldsymbol{n = 2}$ (the first graph), how does the SHAPE of distribution of 500 sample means compare to the shape of the population?	1. Very Different Different A Little Different About the Same	13. Very Different Different A Little Different About the Same	25. Very Different Different A Little Different About the Same
For a sample size of $\boldsymbol{n = 9}$ (the second graph), how does the SHAPE of distribution of 500 sample means compare to the shape of the population?	2. Very Different Different A Little Different About the Same	14. Very Different Different A Little Different About the Same	26. Very Different Different A Little Different About the Same
For a sample size of $\boldsymbol{n = 16}$ (the third graph), how does the SHAPE of distribution of 500 sample means compare to the shape of the population?	3. Very Different Different A Little Different About the Same	15. Very Different Different A Little Different About the Same	27. Very Different Different A Little Different About the Same

The second set of questions is concerned with the **CENTER** of the distribution of sample means.	Joe Normal Dist.	Talia Negative Skew	Cally Irregular Shape
For a sample of size $\boldsymbol{n = 2}$ (the first graph), how does the MEAN of the sample means (**Mean of** $\bar{X}$) compare to the MEAN of the population (the value of μ from the Sampling Distribution Scrapbook)?	4. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	16. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	28. Much Lower A Bit Lower About the Same A Bit Higher Much Higher
For a sample of size $\boldsymbol{n = 9}$ (the second graph), how does the MEAN of the sample means (**Mean of** $\bar{X}$) compare to the MEAN of the population (the value of μ from the Sampling Distribution Scrapbook)?	5. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	17. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	29. Much Lower A Bit Lower About the Same A Bit Higher Much Higher
For a sample of size $\boldsymbol{n = 16}$ (the third graph), how does the MEAN of the sample means (**Mean of** $\bar{X}$) compare to the MEAN of the population (the value of μ from the Sampling Distribution Scrapbook)?	6. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	18. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	30. Much Lower A Bit Lower About the Same A Bit Higher Much Higher

The third set of questions is concerned with the **VARIABILITY** or spread of the distribution of sample means.	Joe Normal Dist.	Talia Negative Skew	Cally Irregular Shape
For a sample of size ***n* = 2** (the first graph), how does the STANDARD DEVIATION of the sample means (**sd of** $\bar{X}$) compare to the STANDARD DEVIATION of the population (the value of μ from the Sampling Distribution Scrapbook)?	7. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	19. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	31. Much Lower A Bit Lower About the Same A Bit Higher Much Higher
For a sample of size ***n* = 9** (the second graph), how does the STANDARD DEVIATION of the sample means (**sd of** $\bar{X}$) compare to the STANDARD DEVIATION of the population (the value of μ from the Sampling Distribution Scrapbook)?	8. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	20. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	32. Much Lower A Bit Lower About the Same A Bit Higher Much Higher
For a sample of size ***n* = 16** (the third graph), how does the STANDARD DEVIATION of the sample means (sd of $\bar{X}$) compare to the STANDARD DEVIATION of the population (the value of μ from the Sampling Distribution Scrapbook)?	9. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	21. Much Lower A Bit Lower About the Same A Bit Higher Much Higher	33. Much Lower A Bit Lower About the Same A Bit Higher Much Higher

The fourth set of questions is concerned with how **SAMPLE SIZE** affects the **VARIABILITY** of the distribution of sample means.	Joe Normal Dist.	Talia Negative Skew	Cally Irregular Shape
Which sample size produced a distribution of sample means with the LARGEST variability (largest value for the **sd of** $\bar{X}$)?	10. $n = 2$ $n = 9$ $n = 16$	22. $n = 2$ $n = 9$ $n = 16$	34. $n = 2$ $n = 9$ $n = 16$
Which sample size produced a distribution of sample means with the SMALLEST variability (smallest value for the **sd of** $\bar{X}$)?	11. $n = 2$ $n = 9$ $n = 16$	23. $n = 2$ $n = 9$ $n = 16$	35. $n = 2$ $n = 9$ $n = 16$
Look at the value for the **sd of** $\bar{X}$ for all three sample sizes. Are any of the values GREATER than the standard deviation for the population (larger than the value of σ)?	12. YES NO	24. YES NO	36. YES NO

8

Active Learning and Technology in the Community College Classroom

Ruby Evans
University of Central Florida

Technological literacy has become the new literacy for the 21st century.

Becoming a teacher of statistics

I spent my formative years in Louisiana, growing up in the Northern portion of the state. My mother, then a librarian, who taught me how to read at age 3, cultivated my interest in education at an early age. My extended family was filled with educators. I began my college studies in the summer of 1976 at Grambling State University (GSU), through a special early enrollment program known as the High Ability Program. GSU, an historically black college and university (HBCU) provided an undeniably invaluable foundation for my personal and professional development. The school's motto, "A place where everybody is somebody," is indelibly etched in my psyche. My professors nurtured my self-confidence and self-esteem.

In 1977, I graduated as valedictorian of my high school class, and continued my college studies at GSU, where I majored in Mathematics Education. While enrolled in undergraduate study, I received numerous awards for academic excellence. Two of the more notable awards were "Outstanding Junior Student, Mathematics and Computer Science, 1980" and the "Award for Excellence in Teacher Education, 1980–81." The Department of Teacher Education at Grambling hosted a formal dinner in my honor to recognize this award of distinction and also presented a plaque to me, certifying that I was the top graduate from the Teacher Education Department in 1981.

I went on to pursue graduate work at three different universities: the University of Arkansas (UA), Fayetteville; Louisiana State University (LSU), Baton Rouge; and the University of Florida (UF), Gainesville. I spent the shortest time in my graduate study at UA. Ironically, that institution had the most far-reaching and long-term impact on my interest in Statistics. In my last semester at UA, summer 1982, I enrolled in STA 5333 Biometry, with Dr. James Dunn. I was intrigued by his classroom presence; by his innovative teaching style; and by his incorporation of technology in the classroom environment. Following this class, I was bitten by the statistics bug and the applicability of the subject matter to daily life. In 1983, I earned a Masters in Applied Statistics from LSU. In 1998, I received my doctorate degree in Higher Education Administration, with a minor in Statistics, from UF.

In 1981, as a graduate teaching assistant at UA, I taught my first college level courses in mathematics and statistics. I believed that students were entitled to quality instruction, an active and stimulating learning experience, and learning-centered instruction. Two decades after my initial teaching experiences, I remain committed to the student-centered philosophy of teaching and learning. I continue to seek ways to encourage students to act, react, and be interested in the teaching and learning process. I continually update my curriculum materials. I set high and consistent standards to both facilitate and maximize student learning. My teaching and learning philosophy derives from decades of student testimonials and narrative feedback about my classroom instruction.

From August 1988 through August 2000, I served as a full-time professor in the Department of Mathematics and Statistics at Santa Fe Community College (SFCC), Gainesville, Florida. My students at SFCC did not live on campus, and as a result, they often used email to contact me. In addition to using the computer to perform data analyses, my students communicated with me online, asking questions about assignments and computer projects. While online, a significant number of my students confessed that they probably would never stop by my office in person, and that the anonymity afforded by cyberspace communication provided a comfort zone for easily approaching me as their professor. Eventually, most of these students found their way to my office for face-to-face conferences.

I required students to cooperate and collaborate on class projects, and to discuss their projects with me on an ongoing basis, prior to classroom presentation. While some students grumbled about the required interaction, these students generally cooperated, albeit not always eagerly, nor readily, with course requirements. Invariably, by the end of the semester I was both exhausted and exhilarated by the symbiotic relationship between the quantity and quality of teaching and learning. The students, whom I taught and believed I "touched" for a lifetime, were also exhausted and exhilarated. Both my students and I learned from each other. I made no apology for caring enough to require that my students interact with me, external to class.

In early 1990, I introduced Minitab in my classes. Until 1995, I was a lone pioneer in integrating the computer into statistics instruction in my department and one of very few professors college-wide, excluding business and computer science faculty, who incorporated any sort of technology into instruction. I was ecstatic when, in 1995, the department offered support by requiring all statistics classes to integrate technology, specifically Minitab. Yet it took almost a full decade for my way of teaching to be accepted among students. For many students, the subject of statistics, especially at the undergraduate level, is perceived as difficult. I believed that if I could get students to adopt a fun, lighthearted, yet disciplined attitude toward learning, success would be within their grasp. By 1998, my students were nominating me for internal and external awards for teaching excellence. By 1999, both the college and my peers were doing the same. It was extremely gratifying to finally have my peers and students join me in embracing the utility of technology.

Despite my commitment to using technology in teaching statistics, I believe it highly improbable that technology can replace faculty. The subject matter of statistics provides an excellent example of why computers and associated technologies are ideal as teacher aids, but not as substitute teachers. Students quickly discover that the computer cannot talk and is adamant in its refusal to explain or interpret the statistical output for them. I view technology as an excellent tool, which assists me in helping my students better learn statistics.

My course

Santa Fe Community College (SFCC), one of Florida's 28 public community colleges, offers 2-year Associate of Arts and Associate of Science degrees, as well as other certificate programs, through day, evening, and weekend classes. The student population is 78.9% white, 9.8% African American, and 5.8% Hispanic. American Indian, Asian, and nonresident alien students comprise the remaining 6.5%. Students between the ages of 15–24 comprise approximately 65% of the total enrollment. The college is a primary feeder institution of upper division transfer students into the State University System via the University of Florida, also located in Gainesville, and "sends more students to neighboring [UF] each year than any other community college" (King, 1998, p. 53).

At SFCC, my principal teaching responsibility was the course STA 2023, Introduction to Statistics, the second largest class in the department. Students who enroll in STA 2023 enter from multiple programs of study and have varying backgrounds with respect to computer skills and technological proficiency. Historically, the majority of students are seeking the Associate of Arts (A.A.) transfer degree. On average, approximately 25 sections of Introduction to Statistics, capped at 30 students per section, are offered in a fall or spring semester at SFCC.

This course introduces students to the fundamental concepts involved in using sample data to make inferences about populations. Included are the study of measures of central tendency and dispersion; finite probability; probability distributions; statistical inferences from large and small samples; linear regression; and correlation.

Day sections of STA 2023 typically follow a MWF sequence. On two of these class days, students meet in a traditional classroom; on the remaining day, students report to one of the department's computer labs for hands-on instruction. If we were scheduled to meet in a department computer lab, I integrated the available technology into course instruction. In randomly assigned pairs at the 15 computer workstations, students viewed an online version of the syllabus, and other course information. For immediate hands-on practice, students were also allowed to experiment with the functionalities of Minitab.

I did provide "chalk and talk" lectures on an infrequent basis. PowerPoint presentations were sometimes used, and handouts of the on-screen slides were distributed to the students. On these handouts, students "jotted" notes as deemed necessary and had the opportunity to participate more actively in the "lecture." The overhead projector and instructor-prepared transparencies also facilitated content presentation and student participation. Student evaluation feedback encouraged retention of these teacher-centered practices, but to a lesser extent than interactive lectures.

During the 14 years that I taught statistics in the community college, the textbook varied. I used texts by authors such as Bluman, Brase and Brase, Freund, Moore and McCabe, Mendenhall, Triola, and Weiss. Regardless of the textbook or course of instruction, I identify a specific set of skills that I would like students to develop in my course. I emphasize student application of critical thinking skills to foster deep learning. I also encourage student to use collaborative learning skills to facilitate "real-world" problem solving. I try to help students develop and/or enhance strong writing skills.

In teaching an introductory statistics class such as STA 2023, my course goals include the following:

- To encourage interactive and ongoing dialogue among students in a learning community.
- To encourage students to be responsible for their education.
- To encourage students to seek out study partners, form study groups.
- To help students improve their college success skills.
- To help students improve their critical thinking skills.
- To help students improve their writing skills across the curriculum.
- To help students overcome their fear of statistics.
- To help students overcome their fears of computers.
- To instill in students a desire to learn.
- To aid students in meeting required computer literacy level for upper division study.
- To sustain meaningful dialogue between teacher and students.

I encouraged students to use the following electronic resources throughout the course: Adobe Acrobat Reader; email; internet access to the web; Lyris list serv; Microsoft Word; Minitab statistical software; Netscape or Internet Explorer browsers; and PowerPoint presentation software.

I asked students to provide a valid email address, and I advised them to check their email daily for class-related information. I also instructed students to check an associated class web site for new uploads in course handouts and materials. The web site provided an online kiosk of practice quizzes, class updates, handouts, and worksheets.

Each semester, I established a single electronic mailing list, based on the email addresses provided by students, for all sections of the course, STA 2023, that I taught. In a fall or spring term, list members for my classes ranged from 90 to 120; in summer terms, from 60 to 90. The mailing list was used, in conjunction with the class web site, to distribute updates and information. Through the electronic mailing list, students frequently queried one another and me regarding content material and application of technology. The ongoing dialogue contributed greatly to an atmosphere of collegiality and camaraderie among students and between students and me. My students and I posted on average in excess of 250 messages per semester to the class list serv. This number excludes the additional private email messages sent to me by individual students.

Early in my teaching career, I began to speak of my classes using a teaching metaphor (Grasha, 1996). My metaphor referred to my classes as flights, or trips, into the air space of learning. As the professor (pilot), I advised students (passengers) to prepare for takeoff on an educational journey (flight). Somewhere along the way, a student came up with the nickname, Captain E, and it stuck. I was known around campus as Captain E and my students nicknamed my sections of the introductory statistics course as "Flight STA2023 with Captain E." I liked the nickname, and encouraged its use by returning papers, sending email, distributing handouts, etc., with Captain E, as my sole identification and signature.

I incorporated the flight analogy into all of my interactions and communications with students: syllabi, classroom presentations, course web sites, electronic communication, face-to-face communication, etc. The metaphor epitomizes my belief that students are entrusted to the care of their professors for an educational journey. It is my responsibility to ensure that students ascend to their maximum flying altitudes, with minimal turbulence. Through our shared flight experience, I feel that I have a leadership role in taking my students where they want to go, to get them safely there, to make their connecting flights. The nickname, Captain E, has transcended both undergraduate and graduate teaching and learning environments.

Over the years that I taught at SFCC, I acquired innumerable evaluation letters and emails from students. Invariably, these messages signaled that students valued the learning communities that we shared. True to my flight metaphor, one student commented: “Like many passengers on Flight STA 2023, I walked on the plane a little nervous and apprehensive about the flight. I knew that my “luggage” of knowledge about the technology was slim but enough to get by. So far the flight has had a little turbulence but nothing that I could not handle. My destination is the University of Florida and my major requires a statistics class before I get there. … I believe that this class is not only teaching me statistics, it is preparing my continuing education and possibly my career.”

A typical class

In teaching introductory statistics, my instructional approach is extremely student-centered, regardless of the learning environment—face to face (F2F), hybrid, or fully web-based. I require that students act, react, and interact with the material, each other and me. Often, this interaction takes the form of a specific handout from a course supplement, a web page, or a specially prepared handout to facilitate an interactive discussion and problem-solving session.

My instructional methodology focuses on my role as a facilitator, rather than a traditional lecturer. My teaching methodology and practices include an online (electronic) student written newsletter, written concept summaries, in/out of class brown bag review activities, required student portfolios, computer problems database, and group research projects, with required classroom presentations. I evaluate student understanding in various ways through cooperative learning strategies and individual assignments. In STA 2023, we either met in a traditional classroom or a computer lab. The students quickly learned the daily procedures by referring to a detailed course timeline, course outline and course syllabus.

In the traditional classroom, I modeled statistical problem-solving through the use of worksheets. For example, I used a worksheet approach to model hypothesis testing. I provide guided practice on a selected problem, leading the students through the process of deciphering which statistical procedure to apply. I encourage the students to remember to look for the key information, just the facts. The students underscore the important statistical information, e.g., mean, standard deviation, sample size, and variance, as they encounter it in each of the problems. Finally, in reply to my verbal cues and prompts, students provide the answers to three key questions: how many populations, what parameter, and what are the problem constraints. This highly interactive discussion method allows students to subsequently determine the appropriate test procedure to use.

At SFCC, both graphing calculators and Minitab were used in course instruction. For me, this was extremely difficult, as I favored application of computer software to handheld technology. Because my colleagues had been reluctant to use computers in mathematics and statistics instruction until 1995, however, most students reported to my classes with TI-83s. The TI-83s, 85s and 89s were used in many of the mathematics classes, so students typically owned one upon subsequent enrollment in Introductory Statistics. On days when we did not meet in the computer lab, students would then solve the problems using their TI-83 calculators.

On days when we met in the computer lab, I modified the worksheet to include problems inclusive of raw data. In the computer lab setting, students would use Minitab to arrive at a solution to the problem. Due to the prevalence of graphing calculators, with which most students had some familiarity, I often allowed the students to use both graphing calculators and Minitab to solve problems, especially in the case of hypothesis testing. With the calculators linked to the computer via the TI-Graph Link cable, students input the requisite statistical data into the TI-83’s. Next, the students perform screen captures of their TI-83 display screens in the Windows operating environment. Students use the TI-Graph Link software to copy the calculator screens to the clipboard. Routinely, students use the mouse to pick and click in sequence — Link, get screen, get screen, clipboard, ok, done — and then paste the screens into a new document in Microsoft Word. Regardless of the scenario—handheld solution in traditional classroom setting, TI-83 graphing calculator solution or Minitab solution in computer lab, I required that students use the following format to write up a brief problem analysis:

- Statement of the problem — Brief summary statement on the problem background.
- Sample statistics and methodology — Identification of the statistical procedure chosen, parameter of interest, problem constraints and summary statistics, and level of significance.
- Statistical inference — Brief comment on the statistical results of analysis, including test statistic, and *p*-value.
- Layman's interpretation — Explanation of the statistical inference in plain and simple English.

Assessment

To measure student learning, I provided multiple assessments throughout the course. These assessments include:

- **Collaborative research projects**
  A required assignment, in which students work in groups of 3–4 to design a statistical study, collect data, analyze the data (via Minitab), provide a written report (Microsoft Word) and present the research findings (PowerPoint presentation) to the class.
- **Computer worksheets**
  A worksheet, which supplements the standard course text, through a hodgepodge of problem types from varying sections/chapters/units of the course. Students are required to use Minitab to complete the assigned problems.
- **Concept Summaries**
  A detailed interpretation, one or more pages long, of a specific statistical topic.
- **Exams**
  Cumulative exams include basic problem solving, as well as open-ended questions for interpretation and stimulation of higher-order thinking.
- **Formal research papers**
  Required, written summaries of the collaborative research projects, in which students write up their research findings, using APA format.
- **Handheld worksheets**
  These worksheets allow students to practice the "plug and chug" mathematics of statistics.
- **Minitab and computer homework problems**
  A database of problems, typically divided into 4 units—Basic Descriptive Statistics, Random Variables and Probability Distributions, Inferential Statistics, Correlation and Regression—that allowed students the opportunity to practice using Minitab in statistical problem solving and interpretation.
- **Portfolios**
  Students prepare summative assessment of all material completed in the course: journal entries, class notes, class handouts, concept summaries, review activities, research project, and documents downloaded from the web
- **Presentations**
  PowerPoint presentations, made by group members, and a required component of the group research projects.
- **Quizzes**
  Periodic opportunities for students to quickly and briefly assess comprehension of material covered.
- **Review activities**
  A fast-paced, highly interactive session, facilitated by the instructor, in which students are given rapid-fire questions to which they must provide answers, either individually or collaboratively.
- **Review for Fun worksheets**
  A fun-filled review of statistical concepts. Students enjoy comparing answers and racing the clock to complete the papers. Bonus points are often assigned based upon completion of the sheets.
- **TI-83 worksheets**
  Hard copy worksheets in which students are required to solve statistical problems, using only the TI-83 graphing calculators. Screen captures were often saved to the computer for later printing and review.
- **Writing Activities**
  Multiple activities, including self-assessments, formative and summative course evaluations (what's working, what's not), journal entries, that encourage students to reflect and write about statistics and the teaching and learning process.

I am always interested in assessing students' perceptions of course instruction. In STA 2023, students appeared to appreciate the multiple pathways to the learning materials. They reported that the lecture supplement, class web site, daily messages through the list serve, in-class group interaction, computer integration, TI-83 graphing calculators, and group projects were invaluable in accommodating varying learning styles.

In my teaching experiences, students consistently identify the existence of a clear course plan, one having a detailed course syllabus, outline, and time line, as being crucial to their overall success. Invariably, students give positive feedback regarding the limited use of "chalk and talk" lectures. Student feedback also has supported the premise that interactive lectures, which incorporate multimedia, are more interesting than the traditional lecture. In the aggregate, students indicate that they can comprehend more information in a shorter amount of time, given these factors.

Students consistently report that these factors—time management, limited computer access, unexpected depth of technology use, and demands of other classes—appear to greatly impact or contribute to failure in a course of this nature. I found that a substantial investment of time and effort, especially for students having little or no computer skills, was necessary for students to succeed in this course. While the workload was heavy for both students and instructor, summary data show that whether or not they were technologically proficient, students who attended class regularly, listened to instructions, and subsequently followed those instructions, earned letter grades of B or better in the course.

Developing an innovative statistics course

I developed my course to reflect standards published by NCTM (1989) and AMATYC (1995). I was also influenced by Annas (1992), who stated that a radical teacher provides a student-centered classroom. In a traditional teacher-centered classroom, instructor-student discussion is limited, and students are relegated to a passive role, with the instructor acting as the sole agent of information distribution. The instructor lectures and students take notes. In developing STA 2023, I openly avoided the teacher-centered course format and opted for the student-centered, active learning environment. My curriculum design involved the creation of multiple pathways through learning materials, to accommodate different student learning styles. In a typical section of my class, students were either required or encouraged to do the following, which are discussed below.

- Participate in interactive lectures.
- Engage in cooperative and collaborative learning activities.
- Produce an end-of-term portfolio.
- Record daily journal entries.
- Read and complete a lecture supplement.
- Write across the curriculum.
- Routinely use computers, handheld technology (e.g., TI-83s), and related application software.

I wanted to aid the students' conceptual understanding of statistical inference; to give students facility with a statistics package for future use; and to create a participatory learning environment by creating active, ongoing, and engaged student involvement. I sought to show my interest in student success by providing review and study aids; prepared lecture notes; lecture supplements; a course website; clear course objectives; formula sheets; help packets; course glossaries; in-class and out-of-class review sessions; and practice quizzes and tests. Eventually, I chaired the textbook adoption committee that chose the first technology-integrated textbook, *Introductory Statistics*, (Weiss, 1995).

I developed a course supplement, "Prepared Lecture Notes in Statistics" which consisted of 13 topics, and a class web site to facilitate the flow of instruction. Through both the lecture supplement and the course website, students were afforded a brief overview of each topic prior to its in-class presentation. I found that when course materials were viewed in advance, class time could be better used to synthesize information rather than merely to obtain facts.

I established a teaching and learning environment in STA 2023 that was aligned with the definition of cooperative learning provided by Hagelgans et al. (1995):

- Students participate in permanent, stable groups.
- A significant portion of the required work of the course is done in groups—in the classroom, in the computer lab, and in homework.
- The evaluation process includes group work.
- A positive *esprit de corps* is fostered among the members of each group.
- The classroom climate is such that a spirit of mutual responsibility develops among group members in the learning process.

I began to use review activities as cooperative learning experiences, with the problem solving being completed by small groups of three to four students seated in proximity. Cooperative learning is the instructional use of small groups so that students cooperate to maximize their own and others' learning (Johnson, D; Johnson, R; & Smith, 1991, p. 12). Informal group activities also allowed the student to identify potential peers with whom they collaborated on a group research project. Students were free to self-select into groups, and many changed groups at least once, prior to deciding upon members for the group project.

Changing a course to incorporate group work takes much thought and effort! An error some faculty members make, particularly in the science, engineering, and mathematics classroom, is not to change assignments traditionally used in the course. Faculty who err in this manner expect students simply to complete the usual assignments in groups. If the traditional problems, experiments, or projects were designed for individual work, they are unlikely to be effective for groups. Moreover, students rapidly perceive that their task could be completed on their own, perhaps even more quickly, and they resent the group work, calling it unnecessary—especially if it is to be done outside of class (Rosser, 1998, pp. 85–86).

The predominant non-residential nature of the community college often contributes to a perception of transient students. The success of the cooperative or collaborative assignment was related to whether or not group projects were authentically shared learning experiences. In my courses, the research projects worked well to facilitate greater understanding among students, increase opportunity for active learning, and provide extended avenues for peer interaction. Those students, having a better grasp of statistical content, reported even greater understanding when explaining concepts and procedures to their peers.

In developing the course, I encouraged students to write across the curriculum. Over the years I used a variety of techniques, including formal research and scholarly papers, concept summaries and a class newsletter, STAT TALK—an online journal which showcased student writings.

I use a summary course portfolio, a research project and an associated PowerPoint presentation, to offer additional ways for students to develop writing, speaking and presentation skills. These activities are also used to help students develop organizational skills.

Future plans

In August 2000, I accepted a position as an Associate Professor of Educational Research at Florida A&M University in Tallahassee, Florida. While at FAMU, I taught graduate courses, primarily in research methodology, Statistics and technology. I also provided web-enhanced instruction (use of Blackboard, WEBCT, web resources, and chat forums). Effective fall 2002, I was named associate professor of Higher Education and Policy Studies, and Coordinator of the Graduate Certificate program in Community College Instruction, in the Department of Educational Research, Technology & Leadership at the University of Central Florida in Orlando. I continue to look for new and innovative ways to make statistics, and any subject matter that I teach, interesting and fun for the students whom I serve.

References

American Mathematical Association of Two-Year Colleges. (1995). *Crossroads in mathematics: Standards for introductory college mathematics before calculus*. Memphis, TN: AMATYC.

Annas, P. (1992). Definition. *Radical Teacher*, 41, p. 46.

Grasha, A. P. (1996). *Teaching with style*. Pittsburgh, PA: Alliance Publishers.

Halgelgans, N. L., Reynolds, B. E., Schwingendorf, K. E., Vidakovic, D., Dubinsky, E., Shahin, M., et al. (Ed.). (1995). *A practical guide to cooperative learning in collegiate mathematics*. MAA Notes 37, p. 8. Washington, DC: The Mathematical Association of America.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1991). Cooperative learning: Increasing college faculty instructional productivity. (ASHE-ERIC Higher Education Rep. No. 4). *What is cooperative learning?* (pp. 1–13). Washington, DC: ASHE-ERIC.

King, E. (1988, November). Community college instructor takes students on a flight in statistics: Integrating software and the Internet expands scope of course. *Scientific Computing and Automation*, 53–54.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: The Council.

Rosser, S. V. (1998). Group work in science, engineering, and mathematics: Consequences of ignoring gender and race. *College Teaching*, 46(3), 82– 88.

Weiss, N. (1995). *Introductory Statistics* (4th ed.). Reading, MA: Addison-Wesley.

Section 3

Statistics Teachers in Departments of Statistics

9

Teaching Interactive Statistics for Understanding

Martha Bilotti-Aliaga
University of Michigan

Students need to talk, debate, disagree, and argue. Class is not chaotic, but it is full of energy.

Becoming a teacher of statistics

I love to teach! I always loved it, and I always will!

When I was 5 years old my kindergarten teacher brought a stool to class, and I climbed on it so that I was able to write on the board and teach my classmates mathematics. That was what my teacher later told me, and that was what I believed.

Although my methods of teaching have developed significantly since that time, my love for mathematics remains intact. The love for mathematics was instilled in me by my father, who always posed interesting and challenging mathematics questions for my two brothers and me to solve. My brothers were very good in mathematics, so I was determined to become very good too.

When I finished high school I had no problem in selecting my career. I applied to study mathematics at the University of Buenos Aires, in Argentina. I was privileged to have wonderful, world-renowned professors who taught me well.

I was learning a great deal and enjoying the university life when many painful personal losses left me devastated. My father died first, then my fiancée, and finally my mother. When I was about to finish my undergraduate degree, a friend, and also the chair of the Mathematics Department, showed me some papers with information about a scholarship to study for a Masters in Statistics at the Centro Interamericano de Enseñanza en Estadística (CIENES) in Santiago, Chile.

“What is statistics?” I asked. I learned enough to motivate me to want to apply for the program. I received a scholarship from the Organization of American States, and I left Argentina to join twenty-one other students representing different countries from Latin America. I thought I would return to Argentina after two years, but I never did return permanently, though I visit often to teach statistics in various settings.

Having been educated as a pure mathematician, I received a shock during the first week of class in Chile when I learned that a good sample size would be 30 or 40, or…. I could not believe that there was not a unique solution. I thought I would not like statistics. However, I was wrong. I soon found that I loved statistics!

Despite two years of intense studying, I have nothing but fond memories of my teachers, classmates, and Chile as a country. I received my Masters’ Degree in Statistics, married my husband who was a classmate of mine from Peru, and was offered a position to teach statistics at the University of Chile in the School of Economics. That is the way I first became a teacher of statistics.

My two years as a teacher of statistics in Chile were very difficult. My mathematical background was strong, and it got in my way by interfering with my ability to create statistical intuition in my students. I tended to depend on formulas and calculations. For me it was natural to look at a formula and to understand its meaning, but the students found statistics difficult to learn and understand. It was my husband who convinced me to teach differently. "Use practical examples," and "Change that theoretical book," he suggested. "Make the students discover!" I tried to incorporate his suggestions and my students welcomed this new teaching method with delight. I had finally grasped some of the right ideas on how to teach statistics.

My husband, our two children and I moved from Chile to the Dominican Republic. There, I held a challenging position at the Catholic University in Santiago de los Caballeros. I continued the work that I started in Chile, namely changing curriculum and working on improving the teaching of statistics. After two years, my husband and I felt ready to continue our educational pursuits. We hoped to enter a PhD program and were accepted to study at the University of Michigan (UM). With two children and a third born in Ann Arbor, I started my doctoral work under the guidance of the best possible advisor, Dr. Michael Woodroofe. My dissertation topic was in Sequential Analysis. In 1986 I got my PhD and became a mathematics and statistics teacher at UM.

In 1998, I became involved in the Calculus reform effort at UM; the chair of the Mathematics Department invited me to participate, along with eight other teachers. Under his leadership, we started to prepare and teach what was called "The New Wave."

The New Wave was the University of Michigan version of the Harvard University Calculus Reform. We adopted the book *Calculus* by Hughes-Hallett, et al. (1998) and with a National Science Foundation (NSF) grant we developed new material, homework exercises, exams, and a program to train graduate student instructors (GSI). The grant permitted us to purchase graphing calculators for each student in the pilot sections.

I loved to teach using the New Wave method, and it was very successful. What I learned from this experience was to use material with real data and real applications and to have students use the graphing calculator to solve real problems. Although I was very reluctant to use the graphing calculator at first, I soon became its advocate and quickly realized its usefulness in the teaching of statistics. As a result, the use of the graphing calculator has greatly impacted the teaching of courses in the UM Statistics Department.

I now use the following principles to structure my teaching, some of which are from the literature and some of which I developed myself:

- The most effective time for learning, retention, and ability to utilize material later is that developmental "moment of readiness." Therefore, that is the time to respond to questions, solve problems, and straighten out mistakes in learning. I will never say, "Just wait until chapter so and so, the answer to your question is there." By that time, that student may not care about the answer any more. This is probably the most serious concern of a new teacher. The questions will not come in a linear form. You have to be very well prepared with a big bag of knowledge to teach using this method. The teacher loses the "control" of the class in terms of what the questions would be.
- The atmosphere to be encouraged is the antithesis of the traditional lecture classroom. Students need to talk, debate, disagree, and argue. Class is not chaotic, but it is full of energy. I encourage students explicitly or implicitly to take risks, to ask questions. I am looking for mistakes. I make it clear all semester, particularly at the beginning, that mistakes are not a reflection of who the students are, what they can or they cannot do, or how intelligent they are. It is only a measure of what must be learned to fix the mistake and increase the understanding!
- Students must be tested on material while still at the beginning of the learning curve, so that errors can be noted immediately, and corrected before they become an erroneous part of the statistical thinking and understanding. Daily questions and frequent quizzes do the trick. When the students finish the quiz and are leaving the classroom, they receive the answer. Seeing the answers confirms either that they were correct or they learn immediately about their mistakes. In tomorrow's quiz that mistake will not be committed again.
- Quizzes assess learning of all previous material, with emphasis on the most current. With this method the final exam becomes another quiz, the easiest one. This is to encourage students to see statistics as a whole, as a method of thinking, and as a method to approach the solution to many problems in life.
- All information about statistics is disseminated in sequences and sizes that mirror students' learning abilities. As their incremental knowledge increases, one can move faster, and also, in larger conceptual blocks. A spiral method of presenting content tests all material again, and again, and presses hard on new ideas.

- The philosophy guiding instruction is that learning is an active problem-solving process. Concepts and procedures are more effectively learned when they are linked with contexts that give them meaning. The emphasis on bringing abstract statistical concepts down to concrete examples in daily life makes statistics exciting and meaningful to the audience. Cooperative learning strategies are used that emphasize collaboration and small-group work (see Lenker 2000, Hagelgans et al., 1995). Hands-on activities as opposed to excessive lecture or individualized seat work are emphasized. A strong emphasis is put on asking questions as well as an expectation that all students should be involved in asking questions (Hewitt and Seymour, 1991).
- Last, but not least, teach with enthusiasm. "The mathematics faculty itself is the most salient advertisement for our field. We should be enthusiastic about what we do, and we should consistently share that enthusiasm with our students. If we exude a high level of job satisfaction we will encourage our students to follow in our footsteps" (see Radunskaya, 2003).

My course

The introductory course I teach, Statistics 100, is an introduction to statistical reasoning. Course topics include methods of analyzing and summarizing data, statistical reasoning as a means of learning from observations (experimental or sample), and techniques for dealing with uncertainties in drawing conclusions from collected data. My challenge is to present statistically sophisticated topics in an accessible, interesting and enticing way to an audience of students lacking a firm knowledge of the vocabulary and symbolic representation. My goal is to open their minds to ideas to help them learn innovative modes of thought that empower them to approach and conquer all sorts of issues within and beyond statistics. I wish to teach students something that they will actually use in their real lives: a way of creating original thoughts.

We usually have about 450 students per semester. The book we use is one I coauthored with my colleague Brenda Gunderson, *Interactive Statistics* (Prentice Hall, 2002). The course is usually divided in four sections. Classes meet three times a week for 50 minutes each, or two times a week for 1 hour and 20 minutes each.

We have four graduate student instructors (GSI) and each of them teaches three labs. The labs meet once a week and students' attendance is mandatory. I usually meet with the GSI weekly for training purposes. We go over the lab, homework problems, and issues that arise that week. If a GSI is teaching the class for the first time, she/he has to sit in a lecture during that semester to learn the dynamics of the course and the way we teach the class.

For technology needs in class, students use the graphing calculator. I believe that once the students visualize a problem with a graph, the problem is easier to solve. Students can be led to many "discoveries" on the spot. The ability to change the setting of the problem and have immediate results and feedback allows the students to think through many new ideas during one class period. These new ideas are timely and relevant and take advantage of the "moment of readiness" to learn.

The use of graphing calculators eliminates much time spent on the manipulations of the formulae, and frees up time for teaching concepts and problem solving. It also alleviates the agony of many of the students who are intimidated by mathematics and calculations. With this technology, every one of my 150 students can have the calculator in his or her hands during the 50-minute lecture. The graphing calculator also accommodates working in groups. Compared to computers, the use of calculators is cost effective and flexible. Most of the students have already purchased the calculator since it is required in the Mathematics Department.

Another tremendous advantage of using the graphing calculator is the immediate feedback to inquiries. Instead of spending time calculating, students spend time thinking, reviewing or expanding the thinking required to solve a problem. It again becomes apparent that there may be more than one way to solve a problem, and there may be more than one answer that is correct. Students use the TI-83 during labs as they do in class.

I start my classes by outlining the topic I will cover, and finish the class with a review of what we have covered. If a student is late, she/he will miss important information. My requirements and information for the class are posted on the web. They include homeworks, articles to read, quizzes and exams (including old exams for review and solutions of the new exams), reviews, due dates, and grading policies. All of these course requirements are explained on the first day of class, including the procedure for calculating the final grade. I found that students want to know the rules clearly and upfront. I prefer it that way too.

During class I use two overhead projectors: one is for the graphing calculator, and the other one is for the projection of transparencies. The transparencies are selected pages from the textbook, but they are in a larger font for easy reading. The transparencies have the page number from the book, so the students do not need to write anything, but instead can open the same page of the book, and concentrate on the solution of the problem and write the solution in the provided space. That way the time is spent on thinking and not copying information which sometimes introduces errors. Because they have enough space to write in their own book, students do not need to bring an additional notebook to class.

When a question needs a written solution, or I need more space to write, I keep the transparency in place but I move to the overhead for the graphing calculator and use this overhead projector to finalize the solution. In that way the students do not need to remember anything as they have the problem and the solution projected on the same large screen that covers the entire wall.

I find that office hours have been completely replaced by email. This email correspondence is good for students and good for me. Instead of office hours, I schedule a weekly, one-hour, review section. I find such reviews to be very useful for the students. I prepare a list of problems from all the previous sections of the book, pass the page to students on a Friday, and ask them to come to the review section with the problems done. Usually the review section is before the weekly quiz. A very wonderful learning environment emerges during the review sections because, as in all my regular class sessions, the students end up finding the solution themselves.

Homework problems are assigned weekly and are graded by the GSI. They are due on Fridays at 2:00 PM and are returned to students during the lab the following week. Every Friday at 3:00 PM the solution of the homework is posted on the web site of the course. That is, an hour after the homework is due, students can check for accuracy of their answers. Homework assignments consist of about 10 problems from the textbook and about four additional problems not from the book. In addition, we give recommended problems that are not graded, but the solutions are also posted with the rest of the required problems.

A typical class

The first class of the semester sets the tone for all other classes. Since cooperative learning and interactive mode is the method I use, I need to form the groups right up front. The room I teach in is usually one of the auditoriums. The number of the students is always 150+. On the first day of class I do not permit entry in one of the two entry doors. With one of my Graduate Students Instructors (GSI), we distribute numbers on the seats as follows:

1	2	5	6	9	10	13	14	
3	4	7	8	11	12	15	16	
18	19	22	23	26	27	30	31	
20	21	24	25	28	29	32	33	etc.

We can prepare this setting in spite of the fact that the auditorium has a slanted floor. Leaving a row in between permits me to reach all of the groups at all times. The room has two columns that section the room into three distinct strata (a wonderful fact that I use when teaching sampling techniques). The numbers on the tables are large enough so that the entering students can spot them immediately.

Another GSI is at the open door with a bag containing the same numbers we distributed on the tables. When everything is ready each student is invited to take a number at random and sit at the table matching that number. I am at the door greeting students one by one. In five minutes students are seated where I want them to be. Otherwise they usually sit in the back and are very unhappy to move to another chair if required to do so. I have learned that asking students to change seats because you want them to be closer to you will take more time, and you will have unhappy people for a while. This is the way I form the groups, and although they are not permanent, human nature leads most of the students to sit in the same seat from that day on.

Next to the numbers on the tables, students will find a page that instructs them to greet each other and fill out two questionnaires, one that is simple to pass to the person seated next to them with: name, phone (optional), email (optional); and another questionnaire with data I will use during the semester. I invite students to talk for five minutes with new acquaintances, telling each other the name of the dorm they live in. This is especially important for working on homework together.

Although the numbers are grouped by fours, I work most of the time with groups of two. It works like this: I pose a question, students think individually, discuss the material with the person next to them, then come back to me with answers for a final class discussion. Sometimes, the students use the other two people seated close to them, but the speed of the class is better designed for a group of two, than for a group of four.

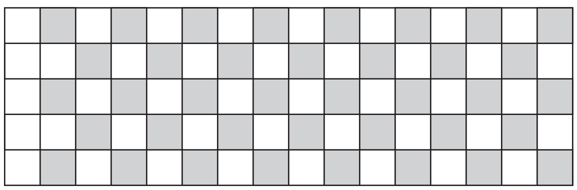
The activities I select to cover in class are the ones that create statistical intuition or make an impact, since the correct answers are contrary to what students expect. I use a lot of graphs and visual aids allowing me to reach all my students, independently of the number of students in the class.

I will describe two of my activities.

Activity #1

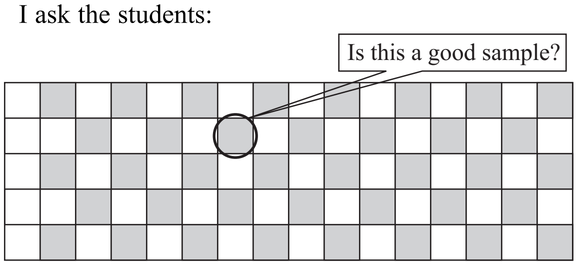
I ask my students: How do you select a "sample size" and "the sample"?

I show them a piece of fabric, for example one that looks like this:



Then I tell the students: "I love this material and think it will make a beautiful blouse for my daughter. But, let me tell you something, when I buy a piece of cloth to give my daughter she usually says, "I love it! Thank you very much." But, guess what? I never see her wearing it. So this time I will be smarter; with your help I will buy a sample so she can see the pattern and really decide whether to buy more or not. This material is very expensive, so please help me to get a useful sample. What should the sample be like? Small enough so I will spend a minimum, but at the same time it should be useful for my daughter to be able to appreciate the pattern.

I ask the students:



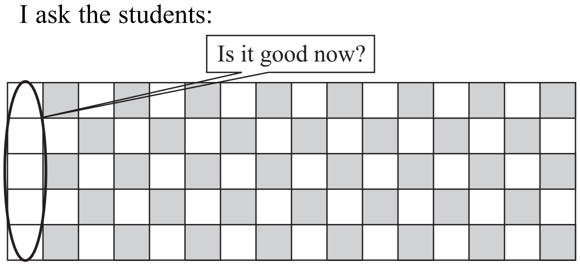
Answer from the students: NO!

I ask the students: Why not?

Answer from the students: Your daughter will not see the pattern.

I say: Mmmm, I see, it is too small! So I will increase the sample size. Now this is a very large, very costly sample.

I ask the students:



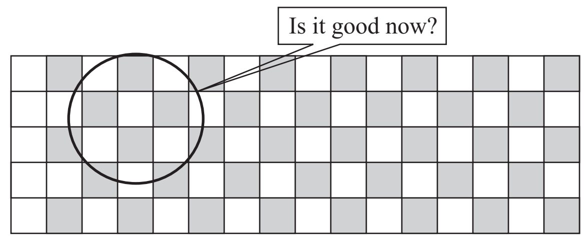
Answer from the students: NO!

I ask the students: Why not?

Answer from the students: Now the sample size is large, but your daughter will not see the pattern.

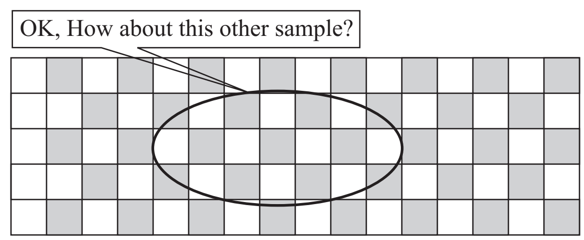
I say: Mmmm. I see, the sample did not capture the variability of the pattern.

I ask the students: Now let's see. Is this OK?



Answer from the students: Yes!, it is small and at the same time it captures the pattern.

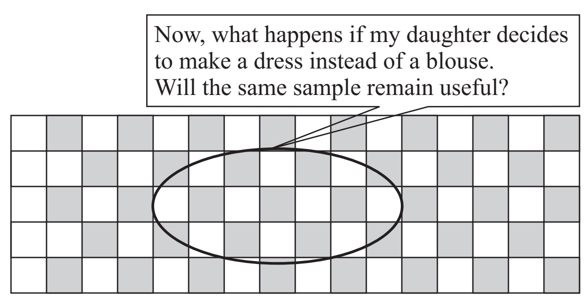
I ask the students:



Answer from the students: Yes! it is small and at the same time it captures the pattern, although the size is not the same as the one before.

I say: So, different sample sizes could be satisfactory as long as they capture the variability of the population. Sometimes you will read $n = 30$, or $n = 40$, or…. All are fine sample sizes. The sample size is not necessarily a fixed number.

I ask the students:



Answer from the students: Yes!

I recap the discovery: For her to look at the pattern for a dress or for a blouse the same sample size and sample will do as long as it captures the variability of the population.

Conclusion: The sample should be small enough to be economical, and it should capture the variability of the population. It does not depend on the population size.

Activity #2

The following are three different data sets. (Note that the three data sets have the same *x*-values.)

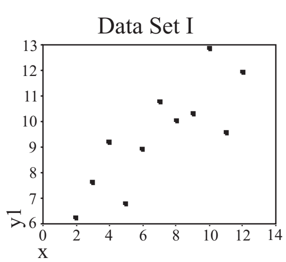
Data Set I		Data Set II		Data Set III	
x	*y*1	*x*	*y*2	*x*	*y*3
8	10.04	8	11.14	8	9.46
6	8.95	6	10.14	6	8.77
11	9.58	11	10.74	11	14.74
7	10.81	7	10.77	7	9.11
9	10.33	9	11.26	9	9.81
12	11.96	12	10.10	12	10.84
4	9.24	4	8.13	4	8.08
2	6.26	2	5.10	2	7.39
10	12.84	10	11.13	10	10.15
5	6.82	5	9.26	5	8.42
3	7.68	3	6.74	3	7.73

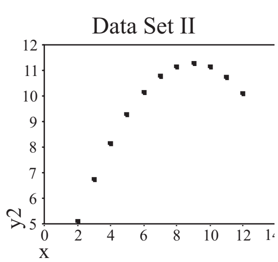
I divide my class into three groups, and I assign one data set to each group. I ask each group to get the linear regression that best fits their data set and to calculate the correction coefficient.

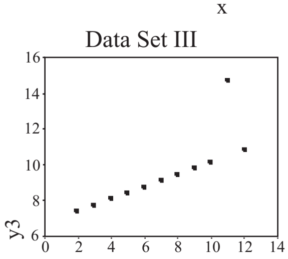
When all the students are ready with their answers, we are surprised that both the linear regression and the coefficient of correlation are the same for the three data sets in spite of the data sets being different:

$$\hat{y} = 0.5x + 6 \quad r = 0.816.$$

I wait for a little while, and I see that students are surprised, but without questions to ask. At that time I remind students of the powerful tool they did not use: Make a scatterplot of the data FIRST! And when they do it, a great surprise occurs because the graphs they see are:







Conclusion: Does it make sense to fit a linear regression line to the data set? Graph first!

Student assessment

Students have six tests during the semester: four quizzes, a Midterm and a Final Exam. Among the four quizzes, I drop the one with the lowest score. Although it seems a regular, traditional assessment, in fact the way the exam is written changes the tone to a modern assessment. The questions assess concepts, some of the questions are open ended, and in other questions, students are required to explain the meaning of any answer given. There are True and False questions, but to answer such questions, students really have to think about new situations.

Another way to assess the students occurs in the interactive teaching during the class period. A question is posted, and the groups have to be ready to defend their answer. This kind of assessment was very useful to me when I was writing the book. When the answer the students gave me was incorrect I asked myself what did I say or not say that made students not understand. If the error came from a reading from the book and I explained the topic in class again, students told me, “You do not say that in the book!” So I corrected the book. I found that if you have good students, most of their incorrect answers came from information incorrectly transmitted.

During the semester, I ask my students to bring the university newspaper *The Michigan Daily* to class, and demonstrate that they know how to evaluate any information given in the published articles and look for ways to improve the written information. In many cases students call the publisher to ask for clarifications. The critiques are given orally in class after I ask students to read a specific article. Students should be able to answer, among other questions:

- Who funded the study?
- Which population is under study?
- What is the sample size?
- What method was used to select the sample?
- Did they check the assumptions to apply a particular method?
- What are the hypotheses to be tested?
- Which hypothesis does the study support?
- Are the graphs confusing?
- Do the tables have titles?

The critiques are not numerically evaluated, but they serve as another way to reinforce my teaching and student learning.

Developing an innovative statistics course

When I was offered the position to teach the introductory statistics class, all the exams were multiple choice and were the only method of assessment used, chalk and board was the method of teaching and only regular homework assignments were given. The examples were not related to real life, but just prefabricated, consisting of a short list of numbers with the task of finding a numerical value of a statistic and plugging values into a formula. No intuition was developed, and no technology was used.

The chair of the department, Dr. Rob Muirhead, gave me a free hand to revise the course and to introduce important innovations. The course developed along with the writing of the text, *Interactive Statistics*, and my chair was very supportive of the writing of the book. At that time I was participating in all possible meetings where calculus reform was developed or implemented and material was being created. It was easy for me to see the parallels to statistics teaching. With Brenda Gunderson, my co-author, we decided to produce a book that would implement ideas of calculus reform into statistics classes. We published the material for use in our class. Some textbook editors saw these materials and came to our office to ask what we were doing. Their enthusiasm and the ongoing institutional support were key factors in the development of our notes into a textbook. I am especially grateful to the Department of Statistics, which insisted that senior faculty sit in my classes and teach Stat. 100 at least for one semester. Their feedback was invaluable. And I am pleased to report that like a drop falling in calm water, the method I use has spread to all other classes in the department.

Future plans

One of my graduate students, Sonya Vartivarian, and I are creating materials for a web enhanced course. Our project is being supported by the University of Michigan Collaboration for Advanced Research and Academic Technologies (CARAT), an Office of the Provost that supports participation in cutting-edge research and educational projects involving advanced networking and computing technologies. The University of Michigan developed what they call a "Living Textbook" that includes a wide range of disciplines, all using the web to enhance teaching and teaching material. Since our web material aligns well with the "Living Textbook," we were asked to add our material and create statistics web enhancements. Only those associated with the university can retrieve this material. To make the material more widely accessible, a CD with our material is now available. The approach of the material parallels the method I use in class, namely the method of discovery and surprise.

References

Aliaga, M. (2001). Information Is Not Knowledge: The Information Revolution. *Bulletin of the International Statistical Institute*. Book 1, 135–137.

Aliaga, M., & Gunderson, B. (2003). *Interactive Statistics* (2nd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

American Statistical Association, Committee on Professional Ethics (1999). *Ethical Guidelines for Statistical Practice*. Can be retrieved from www.amstat.org/profession/ethicalstatistics.html

Anscombe, F.J. (1973). Graphs in Statistical Analysis. *The American Statistician*, 27, 17–21.

Garfield, J.B. (1995). How Students Learn Statistics. *International Statistical Review*, 63(1), 25–34.

Garfield, J.B., & Gal, I. (1999). Assessment and statistical education: Current challenges and directions. *International Statistical Review,* 67(1), 1–12.

Hagelgans, N., Reynolds, B., Schwingendorf, K., Vidakowvic, D., Dubinksy, E., Shahin, M., & Wimbish, J. Jr. (1995). *Practical Guide to Cooperative Learning in Collegiate Mathematics*. MAA Notes 37. Washington, DC: Mathematics Association of America.

Hewitt, N.M., & Seymour, E. (1991). Factors Contributing to High Attrition Rates among Science and Engineering Undergraduate Majors. *Report to the Alfred Sloan Foundation.* Washington, DC.

Hughes-Hallett, D., McCallum, W., Gleason, A., Quinney, D., Flath, D., Osgood, B., Lock, P., Pasquale, A., Gordon, S., Tecosky-Feldman, J., Lomen, D., Thrash, J., Lovelock, D., & Thrash, K. (1998). *Calculus. Single Variable* (2nd ed.). New York, NY: John Wiley & Sons, Inc.

Hutchings, P., & Schulman, L.S. (1999, September/October). The Scholarship of Teaching: new elaborations, new developments, *Change*, 11–15.

Lenker, S. (Ed.) (2000). *Exemplary Programs in Introductory College Mathematics*. MAA Notes 47. Washington, D.C.: Mathematics Association of America.

Radunskaya, A. (2003). A Small Set of Strategies for Increasing Diversity in Graduate Programs in Mathematics. *Mathematics Education Reform Newsletter* Volume 15 (3), 12–13.

Schoem, D., Frankel, L., Zúñica, X., & Lewis, E. (1993). *Multicultural Teaching in the University*. Westport, CT: Praeger Publishers.

10

Teaching Statistics through an Immersive Learning Environment

Deborah Rumsey
Ohio State University

I strive to help students be able to understand, appreciate, use and critique statistics in their everyday lives and in their future careers, regardless of their major, their background, or their prior experiences with mathematics or statistics.

Becoming a teacher of statistics

Ever since my first job as a graduate teaching assistant, teaching mathematics at Illinois State University in 1984, I knew I wanted to teach college students. I felt a connection with students who were struggling, and wanted to help them. I found I was able to help them learn, and be more confident about themselves, and that reward hooked me for life. I found statistics to be a "real world" and vibrant subject, and decided to pursue a PhD in Statistics at the Ohio State University in order to fulfill my dreams of being a college professor.

My teaching philosophy has evolved over the years, and is based on four basic elements: I try to see statistics through my students' eyes; I'm careful about what I teach; I focus on helping students discover the ideas; and I try to keep in mind that attitude is everything. I developed this philosophy during my time in the statistics department at Kansas State University. I still apply the ideas today as I teach courses at Ohio State.

1. I try to see and present statistics through the eyes of my students.

Statistics is a part of all our lives, particularly now in the age of the information explosion. We encounter statistics on TV, the radio, in newspapers and magazines, even in advertisements and junk mail. Politicians, policy-makers, business managers, stockbrokers, doctors, and lawyers use statistics to convey and sometimes manipulate numerical information based on data. Achieving a level of statistical literacy that allows one to receive, evaluate, criticize, and make good decisions about this information is critical for all students, whether they plan to be leaders or participants in this information age. I try to help students develop the following capabilities:

- awareness of where one encounters statistics
- ability to understand a set of basic statistical ideas or tools
- ability to use these tools to evaluate the statistical information being presented
- knowledge of the responsibilities of the government, the media, and the public in the production, communication, and consumption of statistical information.

After trying to teach for many years using the traditional textbooks and methods, I realized I was not focusing on these important objectives, and the connection to real life was all too often missed. In response to this problem, I decided to play a major role in restructuring the introductory level statistics courses for business, physical science, social science, and natural science majors at my former university, Kansas State. I also developed a course for non-quantitative fields called "Statistical Literacy in the Age of Information." This course focuses less on mathematical formulas and calculations, but more on the conceptual ideas of statistics such as variation, sampling, bias, measurement, ethics, and the scientific method, and how they come into play in our everyday lives and in the workplace. Making these changes allowed me to move away from presenting theory first and applications second, to an approach that focused more on interesting questions and scenarios first, involving the statistics on a "what they need to know, when they need to know it" basis. I found the students enjoying statistics more, because they got to see how it helps them answer their questions about the world. I enjoyed teaching more as well.

2. I think carefully about WHAT I teach.

I found myself frustrated at times by the same topics presented in the same order in most statistics textbooks. I started questioning the real need for students to know counting rules and mathematical proofs; and I became discouraged by the sometimes contrived problems that I found myself using in class. I wanted to spend more time on issues like ethics in data collection, how the data were collected (in all chapters, not just in Chapter 1), how the experiments or studies were designed, and how to spot misleading statistics and results. I wondered whether or not I could make these big changes to my syllabus. I was talking with David Moore once about this, and he gave me some advice that has really changed my teaching philosophy: "If we are not going to use a statistical idea or topic later on, or if a statistical idea/topic does not help the students intuitively understand a bigger concept, I do not cover that topic in the class." As part of the restructuring process to align introductory statistics courses with General Education guidelines at Kansas State, I knew then that some difficult syllabus decisions had to be made. After all, introducing activities that instill critical thinking and the ability to connect statistics with other disciplines takes a bit of time away from other topics. Using a new instructional approach which I call immersive learning (telling students less, having them discover more through guided activities) we were able to cover the same big ideas as before, and even added some new topics for emphasis on statistics for every day life (such as the Bonferroni adjustment to avoid the data fishing pitfall). A few topics got eliminated, such as permutations and combinations, and explaining how the binomial probabilities are calculated. A few subjects got moved around so the topics flowed better, using a "big ideas and common threads" approach. For example, the normal distribution was used to figure p-values; the t-distribution came in when we had small sample matched pairs experiments and had to deal with the issue of not having enough data for the normal approximation to work.

3. I tell students less and ask students more questions.

I believe that it is important to present statistical concepts within the context of a relevant problem and let the students discover what tools are required to solve the problem. It so happens that these tools oftentimes involve formulas, but the formulas are not the major focus. If we help lead the students, create opportunities for them to investigate their own questions, and provide support in an active learning, critical thinking environment, they will come up with the ideas necessary to solve the problems. And they will own that process because they used statistics to answer questions of interest to them.

4. I try to remember that attitude is everything.

While it can be argued that this applies to teachers of all courses, I believe this is especially important for statistics. Many students who are not majoring in statistics (particularly those in non-quantitative fields) enter a statistics class with fear and trepidation. They have heard the stories, are unsure about their math backgrounds, and do not understand why this subject is supposed to be important to them.

My objectives in a situation like this are to: 1) hook the student into statistics as fast as possible by showing them where they encounter statistics, and motivating them to become empowered; and 2) work hard to help them develop a sense of confidence in their own abilities as they learn the statistical ideas they need to be good consumers of information.

I try to think about how it felt to be a student, and how difficult a topic statistics was to learn. I devote a great deal of time and effort to including activities, discussions, and examples that come from real life, stimulate their interest and create an active, participatory environment. I found that it helps to carry this philosophy throughout my class, including homework problems and exam questions. I try not to ask students to perform a task just to show they can do it. Every problem I give has a relevant scenario, even a problem as simple as calculating a mean. This really seems to increase motivation, and always presents another opportunity for students to see statistics being used in a positive way.

I have found it helpful as a teacher of statistics to have a personal goal for my students to never have to ask, "Why do I have to learn this?" As I have changed the focus to be more about the interesting questions that statistics can help you solve about the world, introducing statistics through a "big ideas and common threads approach" using a "just in time to use it" approach, I've had that question come up less and less. I'm really proud of that.

My course

The main course described in this chapter is the introductory statistics course for students at Kansas State University (KSU) that I was involved with during my tenure as a faculty member. Jim Higgins and I redeveloped this course, with the help of Lynda Ballou, from 1996-2001. The student audience is basically undergraduate students in the non-quantitative sciences who have to take one statistics course for their majors; we wanted to try to provide some real motivation for them for learning statistics. Our most important goals were to help students become research scientists, and to help students learn the necessary tools, techniques, concepts and ideas from statistics that would help them in their careers and everyday lives— all through the scientific method. We felt it was important to promote statistics as a discipline that is relevant, useful, and important. Showing how it fits within the scientific method, regardless of the science, is a way for students to truly experience statistics at its best.

We had a chance to redesign the course to fit the new KSU general education guidelines in 1995. We started small, by restructuring the syllabus to better meet our needs, outlining what I called the "big ideas and common threads approach." I felt that this was something that typical introductory statistics textbooks and courses were lacking. Once we made the changes, we realized the newly designed course would require more time on the instructor's part in terms of preparation. So we went further, developing a program to train and support the graduate teaching assistants as they taught the course; this led to something that I call "cooperative teaching" where we apply the ideas of cooperative learning to a group of instructors. We adopted the idea of each doing a little, and sharing all our ideas and resources with each other in a cooperative way. Once that was off and running, we took things another step further, reexamining the issue of why students are not as motivated to learn statistics as we'd like them to be— what was missing? This led us back to the roots of why statistics was developed: to help research scientists formulate and answer questions using data as evidence. We realized that the heart of statistics and the key to our approach to teaching it should be the scientific method. So we decided to incorporate the scientific method into our course by immersing students within a context where they get to formulate and answer their own questions using the data they have collected as evidence, and to do that repeatedly throughout the quarter, simulating a true environment where research is being conducted. The context that we chose to meet this end was to have the students explore an unknown planet, called Planet X.

Our newly designed course, Statistics Explorer, incorporated two major elements. First, we wanted to provide students with a relevant, interesting research science environment through exploring Planet X on a series of missions. The missions were designed to emphasize the investigation of students' *own* questions about Planet X, but they were also designed so students would learn the statistical concepts that we wanted to present along the way. Second, we wanted to provide students with a team-based, discovery approach to learning the statistical concepts themselves, through daily activities that were student-centered and teacher/Graduate Teaching Assistant (GTA)-facilitated. In other words, we wanted the students to feel that they were in charge of their own learning and their own investigations, and that we were helping them through that process.

Materials were designed for a computer lab, with a website containing the agenda for each day. We designed a planet that contained one million residents, from different regions/urban centers, and created a database of about 50 variables depicting residents' opinions, characteristics, demographics, etc. The variables were determined using a survey of the students in previous semesters where we asked the question: "If you were going to explore an unknown planet, which items would you be most interested in investigating?" We gave students about 100 items to choose from, and they chose their top 10. Regardless of what major students had, surprisingly they all picked the same set of ten items, including things like what kind of transportation they use, modes of communication, gender, number of children, medical issues, opinions, and sports they participated in. We imposed a structure on the data so there would be various relationships between variables on Planet X that students could uncover in their investigations. We did not want the relationships to be too obvious, and we wanted the relationships to be atypical of what we would find on Earth, thereby forcing students to really make use of the scientific method, and not use any preconceived notions about how things are expected to be. For example, on Planet X, the older you are, the taller you are, and there is no limit to this. We also included an element of no response to the data, and students had to think about how they would deal with that.

We worked with computer programmers on campus to develop an animated sequence where students ship off to the planet, draw their samples, interview some of the participants, and then have the computer collect the rest of the data (to move the process along, once they got a feel for the data collection process). The data from their mission would then be made available to students in spreadsheet form, to be used later for their analysis. Finally, we incorporated MS word into the program, so students could cut/paste results and graphs from Excel straight into their reports at the end of each mission. We then developed the missions themselves, and tied them together in an organized way on a course website.

An online interactive textbook was developed to go along with the course; this book contained the day-to-day activities that students would work through with their teammates to learn the big statistical ideas for each day. It contained all of the exercises, activities, and Excel macros needed to help students practice and discover the concepts. We also included many real-world examples where students could link to an existing website where statistics are being collected and examined in the real world, to be sure that all students know that when they return from Planet X, whatever they learn about statistics and the scientific method can be applied to Earth.

We also designed a special "studio" classroom for this course, to facilitate an immersive, team-based learning environment, and maximize the effectiveness of the teachers who would walk around the room, facilitating the learning. The room has 40 seats, with 5 octagon shaped tables of 8 seats each, and 4 computers at each table. Pairs of students who were randomly paired together share each computer, to facilitate a team learning approach, and eliminate the problem of having a few individual students dominate a teacher's time by asking lots of computer questions. There is no front or back of the room, although there is an instructor podium for brief presentations as needed. Students work together during each class, teach each other, take turns driving the keyboard, and gain confidence and comfort with the technology and their learning process, without the teacher being center stage. Every 3–4 weeks, the teams change, and students evaluate their teammates and themselves on a series of "teammate" criteria. These evaluations compose 10% of a student's grade, to incorporate accountability into the course.

Each class has two graduate teaching assistants (GTAs) as learning facilitators, one acting as the lead GTA and one as the helper. The lead GTA is in charge of what is going on during the class, and makes any decisions. Both GTAs walk around the room, engaging students in discussions, asking and answering questions, and bringing the class together when needed. Each GTA works with two classes per semester, and does the grading for the classes as well. They do not need to put together the daily lessons, but they do have to go over the lessons beforehand, so they can anticipate the questions students will ask.

A typical class

Each day, students enter the classroom, find their seats, and click on the agenda for the day. The first thing they always see is "big ideas for today." This is a list of the 1–3 big points we want them to walk away with. Then, there is a series of activities, each one designed to bring home a big idea. Each activity is designed to be a discovery-based series of questions for each team to work through together. Some activities use links to websites, Excel macros, or graphs that we included for them. Others are just a series of questions meant to lead them through the big idea. For example, for the big idea of being able to connect histograms (visual displays) with summary statistics (standard deviation and mean) we put students in the role of fish hatchery manager, with four ponds of fish. We made four distributions of lengths of fish, one from each pond, and showed the summary statistics for the four ponds without identifying which pond is which (claiming there was a "mix up" in the office, which can happen, as we know). Students have to match things up, and write a report comparing the ponds of fish in terms of their length. Students become quite engaged in this activity and write very colorful reports about it.

After these activities, we have exercises prepared for them to do together to practice and apply the big ideas for the day; whatever work does not get done for the day is taken home to finish. These exercises always included some real world applications of the big ideas learned each day, so students could see an instant connection to the real world. Many of the activities that helped students discover the big ideas also included a real world approach. Quiz and exam questions contained a mixture of both "Planet X" scenarios as well as real world scenarios, applying the ideas learned in class and asking more in-depth questions about the statistical concepts.

Each class period lasted 75 minutes, two times a week. This reduces our start up time, and allows us to make better use of the 150 minutes per week. At the beginning and end of each period, the lead GTA talks for no more than five minutes, summarizing what happened the day before and previewing the new material. They typically also take the last minute or two of each class to close out the day, summarizing the day's activities and looking ahead to the next class period.

Every 3–4 weeks, students are given a "mission" to accomplish on Planet X. The missions are done as part of the daily agenda and are done in stages. Each mission contains four parts, the plan, the data collection (going to the Planet and collecting the data), the analysis (using the statistical ideas learned in the daily activities/exercises) and the report. Each mission involves a specific statistical investigation, requiring specific statistical tools and ideas, but students can choose which aspect of the planet they study on each mission. In terms of the scientific method, students determine their own research questions, collect data on the planet, and analyze their own data and write a report at the end of the mission, all with their teammates. Some of the missions include the following:

Mission 1: Explore the planet by choosing a population of interest, and one quantitative and one qualitative variable of interest, and select a random sample of 25 inhabitants for study. Describe what you find in the data that teaches you more about these individuals, identify a "next question" to ask on a follow-up mission, and discuss the limitations of this study.

Mission 3: Now that you have collected summary data on your population of interest, we want you to go back to the planet and get more accurate information on these same variables. Give us a measure of how accurate your results are, and come up with an estimate of the mean of the population for your quantitative variable.

Mission 5: Select two populations on Planet X and compare them regarding one quantitative variable of your choice. Discuss what conclusions you are able to make regarding your results, and what questions this research leads to next.

The reports students wrote were incredible. It confirms what we already suspected: when students are motivated and buy into the process of their own learning by asking their own questions and collecting their own data, they *can* write, and write very well. Their reports were clean, clear, detailed, and reflected a sense of pride and ownership in the process. They were able to express what they did, what their results were, what the limitations were, and what their conclusions were. Oftentimes, their investigations led them to ask more questions, which were included in their reports. They also worked painstakingly (even beyond our requirements) to incorporate very professional looking charts, graphs, tables, and any results that were obtained using the computer. Everything was labeled clearly in almost all cases. Since students were telling us about their own investigations, it seems to us that they felt more ownership, and hence took more time explaining each step of their investigations, the decisions they made, the problems they faced, etc. Across the board, regardless of the students' major or prior experience with statistics, we found this to be the case.

We found that by having the MS Word and Excel programs available to students in an easy to cut/paste format, students were able to make things look very professional and often spent extra time making sure things looked just right. And students felt more ownership in the process because the facilitators would encourage students to think about what to do next without giving them the answers. For example, they might respond by saying "well, this is your report, what do you think you should do?" or "as long as you can explain why you chose that route, that's fine with me."

We also saw an increased level of accountability in students that we had not noticed before. This may in part be due to the teammate evaluations at the end of each mission. We decided that these evaluations would count 10% of their grade. The evaluation has two parts: 1) evaluate your contribution and performance as a teammate; and 2) evaluate your teammate's contribution and performance as a teammate. Each part of the evaluation contains 5–7 questions, plus an opportunity to explain. Over time, we could see patterns in terms of how people were evaluated by various members of the class, and they usually matched what the instructors had been noticing. We also felt that overall, students were very honest about how they viewed and evaluated their own contribution to the team, as well as their teammate's contribution. We believe that this type of behavior reflects what one expects to see in a real research environment, or in the workplace, and we were pleased with that. One of our goals was to have the classroom experience simulate a research environment, and build students' skills as research scientists.

It was amazing to find out what we *did not* have to tell students. With the right questions and a context where they can use their inquiry skills, students can figure out a great deal on their own. I realized this most profoundly on the day we discussed margin of error. Early on in the pilot phase, before I had an activity written up for this, I prepared a mini-lecture on the topic where I tried to get them to develop the idea of plus or minus about 2 standard errors, in the quantitative case. This was just after we'd talked about summary statistics and the Empirical Rule. I used two fishponds as an example, where length of fish was the variable, the first population of fish had a small standard deviation in their lengths, and the second one had a large standard deviation in lengths. Through a line of inquiry that I used in a 20 minute in-class discussion, students were able to tell me that the margin of error should involve the sample size and the standard deviation of the population. They also told me that Pond 1 would need a smaller sample size than Pond 2 to get the same accuracy, due to the smaller standard deviation in fish lengths. They also realized that as the sample size increased, the margin of error should decrease, and that as the population standard deviation increased, the margin of error should increase. I added a square root to the n, and we had our standard error. Then, I asked how many standard errors should be added and subtracted, in order to have a good idea of the accuracy, and they said "about 2." Now, to us, that was wonderful. It was not important that the 2 should actually be 1.96 at that moment. They understood through the Empirical Rule that this made sense, which was more important. (The 1.96 can be discussed later.) We found that this concept of "about 2 standard errors" can take you a long way in introductory statistics. It really makes it easier to present and discuss the big ideas needed for sampling distributions, confidence intervals, and hypothesis tests, because it's much more intuitive. Students no longer hit that wall right after the Central Limit Theorem, which was a big relief to us!

We also noticed that students were developing a new way of incorporating statistics into their normal everyday language. They were not just saying things the way we taught them to; they were actually using statistical language (correctly) in their discussions with their teammates. We think this is because students spent most of their time in class talking, not just listening. Also, as teams changed every few weeks, the language began to filter throughout this new learning community, and they began to evolve their own way of discussing statistics. We think that through this process, students are more likely to retain statistical concepts because they are able to incorporate the ideas into their own language, although we have not done any long terms studies yet.

Assessment

We assessed the course in a number of ways. First, we designed surveys to assess student satisfaction with the course and compared the results pre and post; we also compared the new approach with the traditional approach, using the same instructors where we could. We also designed exam questions to be common on final exams, comparing the new approach to the traditional approach.

We found that in general students liked the new approach, although at first they were unsure about it. The first time we taught it, students complained that some of the team activities were perceived as "busywork." So we restructured those activities; we also tried harder to explain to students that the process of thinking about the issues, and working problems in small steps without the help of an instructor is the process of learning, not necessarily busywork. And we made sure to be clear about our expectations by including a short presentation on Day 1 outlining the goals for the course and how we planned to achieve them, as well as what we expected the students to get out of the course. This helped make things go much better the next time around.

Student learning was assessed through midterm exams, comprehensive final exams, and daily minute papers/quizzes. Comparisons were made with performance in the same semester of the previous year. Students in the new learning environment averaged 5–10 points higher on exams consistently, and the standard deviation for each of the individual sections decreased to single digits. (Where it used to range from 12–20 it was now around 8–9 points.) Given that there were 8–10 different graduate teaching assistants in charge of their own sections in addition to myself and my co-coordinator, the fact that we were able to reduce the amount of variability in student performance greatly under our new system meant a great deal to us.

Exam questions under the new learning environment were written at a higher level as well, demanding more thinking and reasoning than before. For example, we were able to ask questions like "What do you expect the variability in your sample mean to be if you were to repeat this study again?" vs. "What is the margin of error?" We could ask questions like "Tell us, based on your research, how old you believe the inhabitants of Planet X to be" instead of asking "Give a 95% confidence interval for the population mean age on Planet X." Students consistently performed well on these questions as a group. The best part was that when we asked these questions as research questions, within a relevant context, we did not get the usually disappointing answer where students just report the sample mean. Instead their answers would involve a sample mean *and* a margin of error, along with an interpretation that would be written within the research context as well as a statistical one. In other words, they told us what it meant in terms of Planet X, because they had been there and done that, and they cared about it. Students were also better at determining what type of analysis was needed in order to answer a certain research question. We think this is because through this new learning environment students are better able to understand the scientific method and the role statistics plays within it.

We also asked students whether they thought they learned more in this new learning environment, and we received a positive response on a consistent basis as well. We believe that this positive response is because we focused more on developing the big ideas and the scientific method, offered a relevant investigative context within a team environment, and gave students a chance to explore these ideas in many discussions.

The teamwork component was assessed by conducting teammate evaluations each time a mission ended, as well as an overall survey at the end of the semester. Students liked working in teams for the most part, and they also liked rotating teammates throughout the course; there were a few isolated problems, but they could be solved by having the instructor talk with individuals, or encouraging the teams to work things out. The accountability that came from having students evaluate their teammates as 10% of their grade turned out to be much higher than we ever expected. It was interesting to see how seriously they took this.

We also noticed that student failure/drop rates dropped, and the standard deviation of test scores within sections decreased to single digits. Student attendance was also quite good. We did feel that we were seeing a simulated research/work environment.

The GTAs for the most part enjoyed the experience of working with the course. Most of them signed on to teach later semesters in the new classroom. Many went from being the "helper" in the room to being the leader.

Developing an innovative statistics course

In 1995, Kansas State University adopted General Education (GE) guidelines, stipulating that any GE course contain three components: giving students opportunities for active learning; making connections to other disciplines; and drawing upon their own experiences. Realizing that statistics was a natural fit for these criteria, my colleague Jim Higgins and I set out to restructure our introductory course a bit, to ensure that all sections were being taught in a manner that included these components. We won two seed grants from KSU to develop our course as a general education course.

We came up with a detailed description of what it would mean for our intro stat course to have each of the three components mentioned above. We realized that many teaching resources would be needed for this. We decided that we would take a cooperative teaching approach to this, since the new criteria would involve the graduate teaching assistants (GTAs). We began holding weekly teaching meetings with our 18 GTAs, and started discussing these elements each week, getting them involved in thinking about how they might teach statistics in a way that offered students opportunities to be engaged in their own learning. GTAs took turns giving teaching presentations, testing out ideas, and gaining confidence in methods such as cooperative learning, asking more questions, lecturing less, and learning how to help students learn "what they need to know, when they need to know it." This led to a rethinking of the introductory syllabus, the order in which we presented topics, the manner in which we presented them, and how the textbooks were helping/hurting us in our efforts to make statistics the interesting and engaging subject that we all know it can be.

We realized that students were missing the point in terms of what statistics is truly useful for, which is to help us understand the scientific method. We felt that part of the problem was the fact that we only discussed the scientific method during the first few days of class, and from then on, we basically dealt with problems where someone else had already collected the data.

The other issue was that the traditional way in which we had taught statistics was hindering our abilities to present big ideas in a connected way. For example, we spent two days on the binomial distribution; one day learning how to come up with the combinations for the coefficients, and another day reading tables— all the while knowing that most binomial problems in the real world involve surveys, which use a big enough sample size to warrant using the normal distribution anyway. Some could argue that the binomial is important for small experiments with a yes/no outcome, and I say as long as you can find a way to integrate this topic in with the others in a seamless way, go for it. But you have to also think about how much time the subject will take to explore. I feel that traditional textbooks and teaching methods do not use the "big ideas and common threads" approach. This causes teachers and students to be frustrated by all the topics that must be included in a course. My advice is to have the courage to leave out some of the traditional topics in order to emphasize the big ideas.

Another example is the normal distribution. The traditional approach asks students to pretend that the mean and variance are known for a few class periods, while they work on problems involving the *z*-transformation (and lose an entire day on the *Z* table alone). However, a few days later, students are told that the mean is not usually known, so we estimate using a confidence interval. Then we add in the fact that the variance is not usually known, and hence we estimate it with the sample variance. By the time students get to hypothesis testing, they are confused and frustrated. And while the normal distribution is important of course, we realized that the way we were teaching it (introducing it early on as part of the litany of distributions for their own sake) was not helping students to see the bigger picture. Nor were they able to see how all of this connected to the central limit theorem when the time came. So inevitably, students would fail when we got to those problems where part (a) involved finding the probability that one student scored over 120 on an exam where the mean was 100 and the standard deviation was 20, and part (b) involved finding the probability that the average of 30 students from the class was over 120.

We began to feel that there *had* to be a better way. We decided that what we needed to do was establish the big ideas for each section of material we wanted to present, and find a way to connect those ideas together. This led me to establish what we call the "big ideas and common threads" approach. We took the syllabus and changed it to be what *we* wanted it to be, with a list of three big ideas for each day, in an order that made sense (what you need to know—when you need to know it), connecting the statistical concepts in a seamless way; all within the context of the scientific method. The Empirical Rule took us a long way in our course. The precise values of the normal distribution came into play much later, when discussing *p*-values, so that is where we discussed it.

The syllabus we developed included the following topics in our one semester, 3-credit introductory statistics course (all majors):

- Scientific Method overview (determine research question, plan study, carry out study, organize/summarize data, analyze data, draw conclusions, ask more questions)
- Planning a study: surveys
- Organizing data
  - How?
  - Why?
- Organizing qualitative data
  - Summarizing, evaluating, looking at the distribution, expected value and standard deviation of qualitative data using bar charts
- Organizing quantitative data
  - Numerical Summaries: finding and interpreting center and spread
  - Evaluating numerical summaries found in the media
  - How is the quantitative data distributed? The shape of the data
- How accurate is our sample mean?
  - Sample results vary! What does this depend on?
  - Sample means vary! How are the sample means distributed? What is the mean and standard deviation? Central limit theorem
  - The accuracy of the sample mean: margin of error
- How can we estimate the mean of the population, using the mean of our sample?
  - Basics of forming a confidence interval
  - Confidence interval for the population mean
  - What sample size is needed to get accurate results?
- What is the population proportion? (YES/NO Qualitative Data)?
  - Confidence interval for population proportion (large sample size case)
- Making Decisions when there is uncertainty: overview
  - Variability, probability in decision making, quality control (intro to hypothesis testing)
- Making decisions by testing a hypothesis
  - Converting our sample results to standard units (Z transformation)
  - Measuring strength of evidence: p values (includes reading Z table)
  - Errors that we can make when conducting a hypothesis test
- Comparing two populations
  - Confidence interval/hypothesis test for two population means
  - Confidence interval/hypothesis test for two population proportions
- Drawing Conclusions: observational studies (seen so far) vs. designed experiments
- Planning an experiment
  - Completely randomized designs
  - Matched pairs designs
- Consequences of small sample sizes in experiments: the *T* distribution

- Exploring relationships between variables (quantitative data)
  - Visual displays (scatterplots)
  - Numerical summaries (means, standard deviations, correlation)
  - Drawing conclusions: hypothesis test for the correlation
  - Finding the best fitting line, if appropriate

This syllabus illustrates some subtle changes from a more traditional syllabus. First, there is more emphasis placed on the questions and issues that students might want to ask when they do a study using the scientific method, so the steps in the scientific method are more prominent. Also, items such as the central limit theorem, while important from a statistical standpoint, are not highlighted as the "topic or chapter of the day"; rather, they hold their place on the syllabus as one of the many tools students use to get at the bigger issues. In this case, the central limit theorem is showcased as a tool for helping us look at the accuracy of our sample means, rather than the "crowning jewel of all statistics." We feel this makes the topics more relevant to students and allows students as well as instructors to see the big ideas more clearly, and how they fit within the bigger picture of the scientific method.

Even a slight change in the title of a section, or the language used, can make a difference. For example, instead of talking about *p*-values as a topic, we talk about measuring the "strength of your evidence." This gives the topic more relevance to students, because they will be more likely to want to know how to measure the strength of their evidence, than to just learn about something called a *p*-value. We tried to do this throughout the syllabus.

Once this syllabus took effect, and the GTAs were trained to teach the material this way, we worked on building in the general education criteria: give students hands-on activities, opportunities to make connections, and draw upon their own experiences. This required weekly teaching meetings and strong cooperation among GTAs. We put together a teaching resource notebook of all of our best ideas and shared it amongst ourselves. Each GTA added at least one activity each semester. This helped us a great deal. We also put together exam question pools, where each person submitted a few questions, and they all were organized into the pool, so we could all draw from those as we wrote our exams. Later we went to common exams, but everyone still contributed.

This new system seemed to be working well for students and GTAs. Still, we felt that we could provide an even better context for statistics and its strong connection to the scientific method. We began to wonder: Can we put the entire course within a scientific context or experience for the students, so they are more in charge of their learning? We decided that we could. What is the most exciting context for students to conduct scientific research of their own choosing, yet still learn the necessarily statistical concepts? And how could we serve students from all majors in doing this? Jim Higgins came up with the idea of space exploration, and that fit perfectly. We set it up by telling students that we have been contacted by another planet, and it's their job to explore this planet on a series of missions, each mission geared toward learning more in-depth information about the inhabitants of this planet. We designed all of the investigations so that the students would have to learn certain statistical topics in order to collect or analyze their data.

For example, we knew that we wanted students to learn about how to compare two populations so we designed a mission to do that. We asked students to choose two populations on the planet and one quantitative variable for study, and tell us whether the populations were the same or different regarding this variable. Students by then knew how to collect data on single populations and they tried to apply their ideas to two populations. We had them come up with their mission plan before discussing the sampling issues, so students must think about the real issues involved first. They asked us questions like "Should we take one big sample, and divide the data we get into two groups after that, or should we take two separate samples?" Then we got to respond by asking students, "Well, what would be the advantage of doing it the first way? What are the disadvantages?" "What about the second way?" Then students were ready for our statistical activity for the day, which dealt with sampling issues involved in comparing two populations. They were learning the statistics that *they* needed to know, in order to complete *their* mission. And students came to us wanting to know about it!

I feel that I successfully designed a course where students are able to take charge of their own learning experience, and are able to learn and use statistics to help them in their research investigations. This is how I believe that we should all be teaching statistics—within a proper, relevant context of inquiry—not just for its own sake, as a series of concepts.

Future plans

Since my time at Kansas State as a faculty member, I have moved on to a new position at The Ohio State

University as a Statistics Education specialist in the Department of Statistics. The opportunities that I had at KSU to coordinate courses, manage graduate teaching assistants, and to design classrooms, curricular materials, and learning technology, really helped me to grow as an educator and led me to take on this new, challenging position.

In my current position, I am responsible for providing learning support for thousands of students each year in undergraduate statistics courses. I am continuing my research in statistics education, and am applying my experiences in curricular, classroom, and learning technology design to help students break the vicious cycle of failure that we often see at the undergraduate level when it comes to lower division math courses. I am developing clear, concise and relevant learning resources for students in both statistics and mathematics that I hope will help students feel empowered, and free teachers to be able to explore the "real issues" with their students, rather than have to explain yet again how to use that Z-table, or how to make a pie chart in Excel.

I am interested in the process of how we can make good learning materials, methods, and courses more flexible and customizable so others can more easily and effectively use them. It occurs to me that oftentimes there is one person in a department who has taken a course under his/her wing and totally revamped it with great success, while the rest of the department remains uninvolved and unengaged in the course. To me, this means there will be less of a long-term impact of all this work, if others do not feel that they can or want to pick up where someone else left off, or even want to become involved in a course that is already being coordinated (perhaps too tightly?) by someone else. This includes finding ways to encourage graduate teaching assistants and visiting professors to be able to teach these courses effectively, while still being able to develop their own teaching styles.

11

Becoming an Active Learning Facilitator in the Statistics Classroom

William L. Harkness
Pennsylvania State University

After all those years in the classroom, I think I am now becoming a 'teacher,' although I prefer to think of myself as a facilitator of student learning.

Becoming a teacher of statistics

I have been teaching statistics for 43 years. In view of what I know now, I prefer to forget the first 30 years of my teaching career! For those years there is not much to say, except that my teaching was as traditional as it could have been: lectures three days a week, three or four exams, a few office hours, and get on with the main business of a university (at least according to Deans and Department Heads: do research and get grants). All this changed in 1990.

I was teaching a very large class—about 360 students enrolled—in an auditorium, which appeared to be designed for theatrical productions. The room had a stage at the front, with curtains and lighting equipment, and seats sloping upwards towards the back of the room. The technologies available in it were a blackboard (with chalk) recessed about 15 feet at the rear of the stage and an overhead projector—the choice of nearly everyone who 'taught' in this auditorium. It was the beginning of the third week of the course. I had dimmed the lights (necessary for students to be able to view the screen well) and was about 10 minutes into my lecture when I made a (possibly) fatal error: I looked out towards the sea of faces peering at me. As I glanced over the students in the first and second rows I nodded happily—I recognized most of their faces. But as I looked further out it dawned on me that I recognized virtually no one else, let alone knowing their names.

A wave of anger descended over me at this realization that almost everyone in the class was a stranger to me. I stopped talking. I did not know what to do or say, and started pacing back and forth. Then I noticed that the room had become deadly quiet. It dawned on me that the students perceived something was terribly wrong but had no idea what it was. Thoughts were whirling through my mind—now what should I do? Can I tell them of my thoughts and my anger? I realized I had to do something—anything—to continue. After a few moments of staring at the students and continued pacing, I blurted out 'Damn, I hardly know any of you and that's something I want. If I passed out a sheet of paper for you to sign up for lunches with me, would you do this? At my expense?' The room remained quiet for a few moments and then many students started talking and noisily, perhaps questioning my sincerity at the offer. I said I meant what I offered to do, and began distributing sign-up sheets. I received a list of about 325 signatures that day! Needless to say, I was absolutely stunned by the response. But I did what I proposed to do—I arranged lunches 5 days a week for the rest of the semester, in groups of three to six, and ended up enjoying lunch with about 265 of the students. The cost? It averaged out close to $5 per student, for a total of about $1300.

What lessons did I learn from this experience? For one, I learned that students really wanted to get to know their instructors, and felt good that an instructor wanted to get to know them. I discovered that I was tremendously energized by my interactions with students and they paid much greater attention and respect to me during classes as a result. My daily classes became more engaging and more enjoyable to teach, because I was now teaching people I knew, and not just nameless faces in a large room. In fact, as word got out that I was actually having lunch with students in the class, other nice things started to happen. For example, as I entered the auditorium for classes, descending down the aisle towards the front to the stage, I noticed students were turning towards me and waving to me—it was an experience I will never forget. Boosted by this reaction to taking students to lunch, I continued the practice for the next three semesters—to be sure, the numbers were less but still substantial—usually around 160–180 each semester.

In 1992, another event occurred which drastically altered my approach to teaching. A professor, Barbara Grabowski, called me one day late in the semester (just before Thanksgiving). She wondered if I would give her permission to recruit students in my class to be subjects in a research study. She worked in the Instructional Systems and Design Program. I agreed to help, and to encourage students to participate I offered them extra credit. However, to justify earning extra credit, students were asked to write a half-page to a page description of the experiment they were participating in, and to discuss the randomization of subjects to treatments. When I went to pick up these papers, a few weeks later, we began to talk about my course. I explained that I was now mostly interested in how to get students in large classes more actively involved in learning. A consequence of this interaction was a long-term collaboration with her on research based in my classes. I provided her with a 'lab' (students) for her research and she taught me quite a bit about 'teaching.' She also suggested that there was a much cheaper way to get students involved in learning besides buying them lunches!

In the next few years I learned a lot about incorporating small group activities in the large lecture classes, assigning student projects, and conducting research in statistical education. In 1999 I helped our department procure a very large grant from the Pew Foundation's Center for Academic Transformation to restructure our elementary statistics course. We implemented this restructured course in the fall semester 2000. After all those years in the classroom, I think I am now becoming a 'teacher,' although I prefer to think of myself as a facilitator of student learning.

My course

I have been teaching a class in introductory statistics at Penn State for many years, including every semester for the past 15 years. At the main (University Park) campus, the course is taught in the fall, spring, and summer semesters with an annual enrollment of about 2,200. It is a pre-calculus, introductory survey course. Until recently, it was taught in a 'traditional' fashion (three lectures and two recitations each week). Experienced, full-time faculty lectured to groups of about 240 students. Twelve graduate teaching assistants (GTAs) each taught two, one-hour recitation sections twice a week to about 40 students each. GTAs also held office hours and graded exams. The traditional structure was labor-intensive, requiring many hours of faculty and GTA time per semester, creating resource problems for the department. More important, the traditional structure was not as effective academically as it should have been.

After receiving our grant from the Pew Foundation, we redesigned the course. Our current course has one large group meeting (LGM) per week. The traditional recitation sections were changed to computer-mediated workshops, with technology-based independent learning materials and computerized testing added to give students more practice time and feedback. Instructional roles were shifted from information presentation to learning facilitation. The combination of smaller classes and computer-mediated data workshops enables faculty to have more one-to-one contact with individual students. Faculty are able to address the different needs of individuals, and students can be challenged according to their own skill levels. Computer-based classes enable students to work in teams, which generates more active participation. There is frequent hands-on experience with statistical analysis and with the visualization of concepts. GTA roles shifted under the new structure from instruction to guidance. Technology-based instruction and collaborative activities enable students to explore material that was formerly taught by teaching assistants. Faculty guide many of the recitation section meetings that were formerly led by teaching assistants. GTAs are paired with faculty and an undergraduate intern in the labs, enabling the faculty member to model ways to facilitate learning.

The most important learning goals for the course require students to:

- actively engage with course materials and other students,
- actively participate in data analysis and design,
- understand the reasoning by which findings from sample data can be extended to larger, more general populations,
- understand and apply basic concepts of statistics (e.g., variables, inferences, probability, types of designs, etc.),
- design, conduct, and analyze a scientific research study, including the use of computer software,
- critically evaluate the results of scientific studies.

As noted above, our course consists of two types of classes: Computer Labs and Large Group Meetings (LGMs). The two computer labs and one LGM each week are sequenced as Lab>LGM>Lab or as LGM>Lab>Lab, depending on the time the class is offered. I prefer the first format. We use the recently published text by Utts and Heckard (2001) as the main resource. It has more emphasis on statistical literacy and interpretation than most existing texts.

With just one LGM each week the amount of time available for formal lecturing is minimal. There are about 6–7 lectures in the semester in LGMs. There is some lecturing in labs—up to 10–15 minutes maximum. How do students learn the concepts if we do not lecture? They learn from weekly reading and homework assignments. To motivate students to complete these assignments in a timely fashion, students are given Individual and Group Readiness Assessment Quizzes (RAQs), which are described in detail below.

Students' understanding of the course material is reinforced in the computer labs through individual and group work on activities designed specifically for this purpose. They are given quizzes in about 2/3 of the lab sessions. They are assigned two group projects, the first early in the semester and the second in the last two weeks of the course. Specifically, here is what students are told:

- There will be Individual and Group Readiness Quizzes (RAQs). These will be given in the Large Group Meetings (LGMs). The Individual RAQs will consist of 12–14 multiple choice questions. The Group RAQ will be on the same set of questions and be given immediately after the Individual RAQ. About half of the items will be on previously discussed topics and the other half on new material not previously covered.
- Lectures will be given in the LGMs on days when there are no RAQs.
- In the computer labs you will be working on activities in pairs or small groups to apply what was learned in the readings. Lab quizzes (LQs) consisting of 5–8 questions will also be given in 15–20 of these labs. Items on the Lab Quiz will be based on 1) the activity for the lab and 2) general concepts being illustrated by the activity.
- Some classes and labs will be reserved for work involving the integration of course content, such as evaluating scientific articles and completing group projects.

Three hours of tutorial sessions are set up weekly, to provide assistance to students having difficulties with the concepts and/or with homework. They may also be used to get answers to questions about the reading assignments prior to the RAQs. Tutorials are not designed as lectures or to solve homework problems, but rather to give students feedback on the material.

Our computer labs have PC's, linked to the Internet, which are loaded with Minitab Statistical Computing Software, and on-line quizzing software called 'TestPilot' (2003). In the last five minutes of class students take the short lab quiz on-line using TestPilot. Students may consult with their lab partner or all of their group members (three or four per group) in answering the questions. The purpose of this, of course, is to encourage 'students teaching students'. Student responses on the lab quizzes are sent directly to a file, the results are summarized and made available to instructors to assess student understanding of the concepts covered in the lab. At the next class meeting instructors review any concepts that students did not grasp well. This is great feedback!

A typical class

In the first computer lab I take about 4–5 minutes to look at what we will be doing during the week. This may be followed by a brief overview of the main concept covered in the lab. We have an excellent course management system which provides students with all of the activities we will be doing during the week, including reading and homework assignments, datasets to be used, lab quizzes scheduled, etc. Here is a typical week's page on the web site:

This Week

Week 11: Monday–Friday

Lab 19 Monday

1. *What To Read* Sections 13.1–13.2 before today's lab. Read Sections 12.5–12.7 (covered on RAQ 3) before Wednesday's LGM and13.3–13.5 before Friday.

2. *Exercises* Homework assignment 9: Chapter 12–12.17, 12.21, 12.28, 12.30–12.32, 12.35. Chapter 13: 13.1–13.2, 13.5, 13.6, 13.8. Do at least 6 problems from Chapter 12 and 4 from Chapter 13.

3. Today's lab activities will be concerned with inference for one population mean, including paired comparisons: confidence intervals and testing.

4. *Datasets* Use the data from the **Spring Survey**.

5. *Study Guides* RAQ 3 is scheduled for Wednesday's LGM. Take a look at the Study Guide for Chapters 9, 10, 12 and 13.

6. *Test Pilot* Take the Lab Quiz.

LGM Wednesday

1. RAQ 3 will be given today. Please bring a #2 pencil. Coverage for the RAQ is Sections 9.3–9.7, Chapter 12, and Sections 13.1–1.2

Lab 20, Friday

1. We need to form groups for Project II. Please login on to coursetalk (an asynchronous web-based communication software, like a chat room) and check your membership. If you are not listed, contact Yudan or Eliza (GTAs) to correct the situation. If you want to change groups, we can consider that as well. The URL is coursetalk.cac.psu.edu

2. *Activities* Today's activities will be on two proportions and two means. We will look at whether or not the proportion of students who would have sex without being in a committed relationship is the same or not for whose who have driven under the influence and those who have not. In comparing two means we will get confidence intervals and tests.

3. *Datasets* Use the data from the Spring Survey

4. Use the *t* -table* of multipliers

5. One of the questions on the Lab Quiz asks you what type of general 'theme' is most appealing to you. Confer with members of your group to express your preference.

6. *Test Pilot* Take the Lab Quiz.

The schedule for Monday's lab is fairly typical. After the brief look at the week's agenda, students begin working. As soon as they enter the lab, they login on their PC, get on the internet, bring up the week's agenda, and then open Minitab. They return to their 'This Week' page, open 'Datasets' to select 'Spring Survey', then copy and paste this dataset into their Minitab Worksheet. 'Spring Survey' is a file of data collected from the students during the first week of classes. It consists of about 40 items on a variety of variables, such as gender, grade point average, race, eye color, height and weight, ideal height and weight, drinking habits, etc. These data are then used throughout the semester to illustrate statistical concepts learned in class. A student intern hired by the department passes out an 'activity' for the students to work on, while a teaching assistant at the front of the lab does the same thing as students are to do, but on a delayed basis. I walk around the room prepared to help any students having problems or to answer questions.

After the students have worked (individually, in pairs or in groups of four) about twenty-five minutes on their activity, I ask for their attention again and we discuss the activity. Here is an example of a short activity we use to learn about the sample standard deviation and the empirical rule: (Note: we have already studied confidence intervals):

Activity #013
Age Measurements for the Shroud of Turin

The Shroud of Turin is a linen fabric that, since 1354, has been claimed to be the burial garment of Jesus Christ. In efforts to establish its authenticity, there has been an enormous amount of scientific testing performed on this object. In one study, several small strips were sent to labs in order to perform radiocarbon dating, a process by which the age of items can be estimated (with some degree of uncertainty). Four of the strips were sent to a lab in Arizona in 1988, resulting in the following estimates for their date of origin: 1397, 1298, 1382, 1287 AD, and their dated age in 1988

date of origin	age in 1988	$(x - \bar{x})$	$(x - \bar{x})^2$
1397	591		
1298	690		
1382	606		
1287	701		
$\Sigma x =$			$\Sigma(x - \bar{x})^2 =$

Use the above table to guide you through the hand calculation of the standard deviation.

Using the empirical rule, does it seem likely that the fabric could be from the time of Jesus Christ's death?

In Minitab, double-check the standard deviation that you calculated. To do this, enter the four data points in a column, then obtain the descriptive statistics (as you've done before).

Now find a 99% confidence interval for the population mean year of origin for the strips sent to the Arizona lab. Go to Stat > Basic Statistics > 1-Sample *t*. In the "Variables" area, select the column into which you entered the data, then make sure that the confidence interval level is 99. Interpret the range you obtain. What do the results suggest about the authenticity of the shroud (based on the strips sent to the Arizona lab)?

If you think it would probably have been a good idea to send strips to other labs for dating, you are right—this was actually done. If you are interested in these and other data and information, the web has many sites dedicated to studying the Shroud of Turin, such as *The Shroud of Turin Website*, *The Shroud of Turin: Genuine artifact or manufactured relic?* (by Jack Kilmon), and *The Council for Study of the Shroud of Turin.* You may find the controversial debates on these sites to be interesting.

About 15–20 minutes before the end of class, I get the students attention and we discuss the work they have done. This includes going over their solutions and interpretation of the results. It also provides students with valuable feedback on the concepts they have been working on and their understanding of them. In the last 6–8 minutes of the class they take the on-line lab quiz covering the day's concepts and the activity they were working on. We found that by putting questions on the lab quiz about the activity that students were more attentive to their work.

As for the individual and group collaborative activities, we were pleasantly surprised at the students' positive reactions to not being lectured to and instead being able to work in groups in the labs to apply what they had learned from the resource.

Assessment

We use a variety of assessments in our introductory course, including: in-class exams, a final exam, Readiness Assessment Tests, homework assignments, projects, and lab quizzes.

The two in-class exams are 50 minutes in length—open-ended questions count for about one-third of the grade and multiple choice items about two-thirds. GTAs grade the open-ended questions and the multiple choice items are machine-graded—answers are placed on bubble sheets (Scantron Forms). The multiple choice portion is graded immediately (usually same day as the exam), scores emailed to students, and an item analysis performed showing (i) student responses, (ii) % correct for each item, (iii) a test validity measure, and (iv) average and standard deviation for all sections of the class. Instructors use the item analysis as an assessment instrument for identifying concepts that students did not grasp well.

I normally assign 12 weekly homework assignments along with the reading assignments. Some of the homework is discussed in labs or in the LGM. Undergraduate students are hired to grade these assignments, under the supervision of GTAs.

The final exam is usually entirely multiple choice because of a requirement to assemble and report grades to the registrar's office in a very short time period (two days or less). If the exam is scheduled early in the finals week, a portion of the exam (up to 50%) may be open-ended.

The Readiness Assessment Quizzes (RAQs) have two major components, and sometimes a third, depending on an instructor's use of them: (i) an individual component, (ii) a group component, and (iii) an appeal process. RAQS were developed by Michaelson (1999) as an instructional and assessment tool. Students are given reading assignments before classes and prior to instruction on the material.

The goal of these reading assignments is to provide motivation for students to learn some of the basic concepts of the course on their own. After doing the reading assignments students come to class and take a RAQ, made up of true/false and multiple choice questions. These questions test knowledge and understanding of general concepts and principles rather than small, detailed facts. The goal of the individual RAQ is to ensure accountability and understanding of the reading assignments. Students take the individual RAQ first and turn it in and then immediately re-take the same test as a group (previously set up) of three to five. The goal of the group RAQ is to foster students helping one another comprehend ideas that they may not have gotten on their own. If the instructor chooses to do so, students are allowed to appeal any incorrect answers based on the quality of the question or a justification for their answer choice. Each student receives an individual and group grade for each RAQ.

The instructor uses the feedback from the individual and group RAQ scores to determine where students still have misconceptions or misunderstandings. The concepts that students did not get on their own can be used to guide and inform instruction. The feedback helps the instructor focus instruction and activities on application of the course content rather than spending time covering concepts students can easily obtain through self-directed reading and learning. Course activities are typically completed in pairs or groups. The RAQs and the content covered on them are used as a means to prepare students for the application of the content in problem-based activities.

RAQs cover natural units or modules, usually one or two chapters in the text. Each semester about five or six

RAQs are given. RAQs provide a powerful motivator for students to read material prior to classes (since it is a major component of their grade) and to keep up with work on a regular basis rather than trying to study at the last minute before an exam.

We believe that the assessment of student understanding of concepts using RAQs has proven to be very effective in detecting areas in which students are not grasping the concepts, thereby enabling corrective actions to be taken in a timely manner, and in preparing students for higher level activities in the computer labs than previously. As a result, students have been helped in building skills, as the evidence of pre- and post-tests show. The web page enables more rapid feedback to students, another crucial element in the learning process.

Student perception of the importance of RAQs is evident in the results from Innovation and Quality (IQ) survey data where the majority of students (55%) rated the RAQs as one of the most important aspects of the class. Seventy-five percent of respondents believed that periodic RAQs help them keep up with the readings and that they were vital for their learning and understanding of the content. As voiced in focus groups, students felt that the RAQs helped by promoting recognition of holes in their understanding. In addition, students liked the opportunity to work in groups and interact with others in the class. Most students emphatically suggested keeping the Readiness Assessment Quizzes as part of the course.

Projects. Students are assigned two projects, the first about the fourth week of the semester and the second during the last two weeks of the semester. The first project is moderately well-structured and the second is fairly unstructured, with general guidelines provided.

For the second project, a survey is developed cooperatively by the students and me. Students are given a scenario, such as:

> A President of a music company has experienced a downturn in his business. He asked the marketing department to collect data on aspects related to music, and the Director of Marketing contacted Penn State to do a survey of college students for them. Naturally the Department of Statistics was contacted to do this. So, with the help of the students a survey is created. As a group project, students are asked to analyze the data, interpret it, and prepare a report for the Director of Marketing at the company.

The requirements imposed on the project specify that students are to perform at least six different statistical techniques in their analysis (e.g., compare two means, two proportions, regression, chi-square test, analysis of variance, etc.). All members of the group initially receive the same grade. But adjustments (upwards and downwards) to their grade are made based on the group member's assessment of each group member's effort and contributions to the successful completion of the project. This is done on a survey form prepared specifically for this purpose. A 'slacker' can get a 'zero' if s(he) did nothing in the judgment of the other group members and an especially good effort as expressed by the group can lead to a reward through an increase in one's grade.

Developing an innovative statistics course

As a consequence of my involvement with my colleague in Instructional Design and my efforts to engage students in my large lecture class, I was invited to become a Faculty Fellow of the Schreyer Institute for Innovation in Learning at Penn State. Gradually other instructors in the department decided to try some of the methods I was using to engage students. Soon there was a cadre of six or seven of us meeting weekly. We called ourselves TAPS for "Teachers of Applied Probability" and "Teachers at Penn State." In 1998 we submitted a group proposal for funding from the Schreyer Institute, which was accepted. This was the first time an award had gone to multiple submitters and it was viewed as a departmental proposal. We became a focal point of the university's effort to get departments to change the way they taught. This in turn led to our being considered as the university's candidate to submit a proposal to the Pew Foundation, in a competition with other departments and programs. We subsequently received the Pew Grant. In setting out to restructure our course, we met as a group for about two months discussing what we wanted to do before we began working seriously on the redesign. We arrived at a point where we identified four specific academic problems to be solved:

- The traditional format did not address the broad range of differences in student learning styles and quantitative skills.
- Students with weak skills need more individual attention and more opportunity for group collaboration, while students with strong skills could benefit from having more opportunity to explore the material more fully.
- The original format did not encourage active participation. It was difficult for students to ask questions, discuss the material, or collaborate with other students. There was not enough hands-on experience with data analysis and collection.

- The traditional structure required twelve GTAs each semester. It was difficult for the department to identify, much less allot, this many qualified assistants for this course. Most graduate students in statistics have undergraduate degrees in mathematics or a scientific discipline, limiting the effectiveness of the statistics instruction they can provide. As initially designed, the course did not provide tutoring assistance for students. GTAs working with a particular faculty member had only two or three office hours for students within their own recitation section, so students received little individual attention.

We had additional concerns: students were unable to apply statistics in follow-up courses, had a negative attitude towards statistics, and retention of subject matter appeared to be very short.

We strongly believed that there was a need for change in the traditional learning environment to enable students with a variety of learning styles to engage in classroom discussion, collaborate with fellow students, and participate in one-to-one interchanges with trained instructors. We asked "What should we be teaching in the beginning statistics course?" and "When should we do it?", "Where should we do it?" and "How should we do it?" We decided that we wanted to set up an alternative assessment process, put emphasis on hands-on learning and group work in computer labs, and re-evaluate sequencing of topics.

We made several learning assumptions: learning is enhanced by collaborative group activities, students can learn independently, and, of course, there are different types of learners. We did a lot of reading about learning groups. We looked at technologies and assumed that (a) computer labs dedicated to instructional use would be available, (b) statistical software would be available online in the labs, and (c) instructors and students can communicate synchronously and asynchronously. Finally, we adopted some guiding principles:

- give responsibility for learning basic concepts to students and let the instructor's role be that of a facilitator,
- provide as much hands-on practice as feasible,
- use technology appropriately,
- reduce time between concepts and applications to a minimum.

As for sequencing, we attempted to free ourselves of pre-conceived notions on the order in which they are covered. We then reassembled a sequence of topics that we felt met our course goals. We decided that we wanted to:

- introduce inference very early, within the first two weeks,
- include inference with new topics as we introduced new topics—for example, in covering descriptive statistics, use case studies, formulate research, null and alternative hypotheses, obtain *p*-values (initially without showing how they were obtained, leaving that for later),
- expose students early to decision-making and interpretation of analyses.

As a result of this process we thoroughly restructured the course, as described above. We placed less emphasis on formulas (but reaffirm their importance in summarizing concepts, like the one sample t-test statistic involving a sample statistic, a null hypothesis, a standard error, and a sampling distribution).

To determine the impact of our revised course on student learning, a content knowledge test consisting of 18 items was developed prior to the restructuring of our introductory statistics course. It was administered at the beginning and end of the spring 2000, fall 2000 and spring 2001 semesters. During the spring 2000, two sections were taught in the traditional format ($n=340$) while one section was taught as a pilot using the revised approach. In the fall and spring semesters of 2000/01 all classes were taught using the new format.

A 20-item test on choosing the appropriate statistical technique from a set of 10 was created and has been used as part of final exams every semester since then.

Statistics on the number of students who received Ds, Fs, student grade point averages, and number of course dropouts were compiled for the 5-year period 1996/97 through 2001/02.

Assessment of student performance in follow-up courses using a subset of the content knowledge test is on-going.

The results of all these types of evaluation revealed many positive results. The pilot and redesigned classes outperformed the traditional class on the final test of content mastery by 10% to 13% (60%: traditional class, 66% in the pilot class, 68% in the redesigned classes). The improvement in performance in the redesigned class was greatest on concepts. On technical aspects (working with formulas, for example) the traditional class performed slightly better. Students in the restructured course were able to identify the correct statistical technique to use about 86.5% of the time, about 11% better than the 78% correct rate for students in the traditional course. This is viewed as a consequence of lab work. The percentage of students receiving a D, F, or dropping the course decreased from a rate of about 12% in the traditional course to about 9.8% in the restructured course, or about

18% decline. The average student grade point average was essentially unchanged: 2.974 in the traditional course and 3.015 in the new course.

Follow-up data on the performance of students in subsequent courses is too preliminary to draw any conclusions. We will continue to monitor follow-up performance to see if the pattern shown above will be applicable to this as well. One final important evaluation result: our course redesign is saving the department about $115,000 per year!

Future plans

We submitted a proposal to the National Science Foundation to restructure our introductory biostatistics and engineering statistics courses and it was approved. I will be working intensively on this during future semesters. We are so happy with our current redesigned elementary course that we wanted to explore using the model in other first-level courses.

Penn State University has about 20 other campuses around the state (in addition to the main campus at University Park). We have been cooperating with instructors at these campuses in developing a modified version of our model that fits their needs.

We are constantly making revisions in our restructured course—but the basic structure has been maintained. We consciously built in procedures for change to encourage modifications—we did not believe we had developed the perfect course (whatever that might be!) and did not want to discourage innovation in the future.

References

Felder, R. M. and R. Brent (1996). Navigating the bumpy road to student-centered instruction *College Teaching*, 44(2), 43–47.

Michaelson, L.K. (1999) Myths and methods in successful small group work. *The National Teaching and Learning Forum*, 8(6), 1–7.

Michaelson, L.K., Black, R.H., and L.D. Fink (1996). What every faculty developer needs to know about learning groups. In L. Richlin (Ed.), *To Improve the Academy: Resources for Faculty, Instructional and Organizational Development*. Stillwater, OK: New Forums Press.

Springer, L., Stanne, M.E., and S.S. Donovan (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69(1), 21–51

Test Pilot (2003). Retrieved October 16, 2003, from www.clearlearning.com

Stat 200 website (2003). Retrieved October 16 from stat200.cac.psu.edu/

Utts, J. M. and R.F. Heckard (2001). *Mind on Statistics*. Pacific Grove, CA: Duxbury Press.

12

Integrating Pedagogies to Teach Statistics

Beth L. Chance
California Polytechnic State University, San Luis Obispo

Nothing is more exciting than hearing students relate something from my course to their other life experiences or to see them finally appreciate how statistical evidence allows us to make tentative conclusions. It is also gratifying to observe students simply realizing that this is material they can actually learn and master.

Becoming a teacher of statistics

As a math geek my entire life, I was always drawn to the applications of mathematics, to working with numbers to solve problems, whether it was traffic flow or cancer research. But I was also frequently frustrated that my friends did not share my joy for mathematics. In high school especially, I often wondered if their appreciation for the subject could be increased through different teaching methods. I found I truly enjoyed helping my friends understand something and became very interested not only in playing with numbers myself, but also in trying to find different ways to explain mathematical ideas to people.

As an undergraduate mathematics major I did not have a very data-oriented introduction to Probability and Statistics, but did enjoy the problem-solving approach in my Operations Research (OR) courses. At the same time, I was also drawn to Psychology classes. I earned a minor in Psychology, worked on a research project for a Cognitive Psychology instructor and read about cooperative learning, constructivism, computer-based diagnostic systems, and "meaningful learning." Initially I was not sure how to combine my loves of Mathematics and Psychology into a career, and was torn between pursuing graduate study in OR or in Educational Psychology, even after visiting schools in both areas. I believed at the time I would be keeping my options open longer (and my earning potential higher) by pursing an OR degree. I enrolled in a rather prestigious but theoretical doctoral program in Operations Research at Cornell University, and again gravitated toward statistics and data analysis. I also took advantage of Cornell's flexibility and designed a minor in Education, including courses in Mathematics Education, Science Education, and Curricular Design.

At Cornell I served as a teaching assistant for several courses. This included leading recitation sections and grading homework assignments. I loved the experience and believed I was doing a good job based on student response. As graduation neared, I eagerly pursued teaching positions, but again was a little unsure where I belonged. As I wanted to focus on teaching, I did not even look at research universities. I was also interested in doing research on teaching but was not interested in trying to teach in a mathematics education program. I ended up at the small, teaching-oriented University of the Pacific (UOP), a private university in northern California.

At UOP I quickly learned that I had much more fun as the TA than the teacher. As the teacher, I could not blame the students' frustration on anyone other than myself. I also realized that my background in Statistics was very different than the course I was asked to teach. I feel I learned more about Statistics from reading Moore's *Concepts and Controversies* (2001) prior to teaching my first course than I had in my entire mathematics degree programs. Designing lectures and meeting my own need to learn the material before teaching it led to many late evenings.

During my first year at UOP, I was fortunate to attend workshops for the *Activity-Based Statistics* materials (Scheaffer et al, 1996) and for the CHANCE course at Dartmouth. I used activities from these workshops in my classes and saw an immediate impact on students' engagement. These workshops were also influential to me in that I was able to discuss teaching ideas with other instructors in similar situations (and realize I was not alone) and I developed very valuable professional contacts and resources.

The next year I began designing a lab component for my course and added a one-hour/week meeting in a computer lab to the course structure. With a small university grant I put together a lab manual, stealing ideas from sources such as *Journal of Statistics Education*, the *Elementary Statistics Laboratory Manual* (Spurrier et al, 1995), and other instructor lab manuals (Egge et al, 1995). I began giving talks at meetings such as the Joint Statistical Meetings and the Joint Mathematical Meetings on some of these implementations and I was hooked. This began my involvement in the statistics education community.

I feel very fortunate that I began my teaching career at a time when there was much interest in how statistics was being taught and in recent recommendations for the reform of statistics teaching. I found myself in complete agreement with many of the suggestions for teaching statistics: the need to engage students, to focus on real data, and to emphasize concepts rather than computations. Statistics, Probability, and Data Analysis are continuing to permeate our everyday world as well as the elementary and secondary mathematics curriculums. I feel it is very important that we develop quantitative literacy for all students, from elementary school on, and statistics provides a direct connection between the mathematics students need to know and the quantitative reasoning they will need to employ.

I now teach in the Statistics Department at Cal Poly, San Luis Obispo (a public university with approximately 17,000 students). Many of the students I teach at Cal Poly are freshmen in fields such as psychology, journalism, recreational administration, agricultural business, and education, to name a few. These students are enjoyable to work with and often bring diverse experiences that enrich the classroom. They often do not understand the relevance of statistics to their programs of study and sometimes need prodding to take advantage of the technology available and to invest quality time in thinking about the material. But there is nothing more exciting than hearing students relate something from my course to their other life experiences or to see them finally appreciate how statistical evidence allows us to make tentative conclusions. It is also gratifying to observe students simply realizing that this is material they can actually learn and master.

My favorite teaching moments are when I am wandering around my classrooms where discussions and debates about statistics fill the air from students who are truly engaged in what they are learning. I am not sure I personally could accomplish this level of engagement teaching another subject.

My course

Most of the following discussion will concern the "Math 37" course that I taught for five years at the University of the Pacific. This general education (GE) course was intended primarily for science and business majors, as we had a different course for social science and education majors, and another for engineering and mathematics majors. The prerequisite for this one semester, four-credit hour course was intermediate algebra. Most students would take the course to fulfill the GE requirement or as required by their major. These students were not expected to become practitioners of statistics, but definitely needed to know how to read and interpret statistical findings. My goals in teaching this course were to enable students to

- Interpret and critique published statistical results and communicate using the language of statistics.
- Understand and employ proper data collection strategies.
- Understand key statistical concepts such as randomization, graph-sense, resistance, association, chance, sampling distributions, confidence, significance.
- Be wary of the cautions and limitations of statistical analyses.
- Read and discuss computer output, as well as utilize a common statistical package to analyze data.
- Gain experience working in collaborative teams with diverse student backgrounds.
- Increase their appreciation for the power of Statistics and their understanding of the overall statistical process.

After moving to Cal Poly I developed several courses in this environment. While my goals have remained consistent, some of my discussion will include adaptations I have made in teaching on the quarter system, with all class meetings occurring in a computer classroom. While these courses are not as developed as my UOP course, I hope discussion of this transition will also be of value.

At UOP, classes of 35 students met three days a week for 80 minutes in a traditional classroom. Students entering the classroom picked up a lecture handout, typically one or two two-sided pages that contained details of the examples we would discuss and summaries of important results, with space left on the handout for students to fill in calculations and explanations discussed in class. (These handouts were also available to students from a course webpage. Examples can still be viewed at www.rossmanchance.com/uop/chance/math37.html#Lect.) Students also saw a list of topics from the previous class meeting summarized on the board. They were encouraged to write this outline down before we discussed it briefly at the start of the period, and they were encouraged to ask questions on this earlier material.

Most of my presentation was via overhead transparencies with empty spaces similar to those on the student handout that I would fill in along with the students as they provided me the answers. The content of the lectures followed fairly closely the text, *Introduction to the Practice of Statistics*, third edition (Moore and McCabe, 1999), though I covered the chapters in a different order. Most topics were motivated by an introductory example, and then we completed several follow-up examples, often using data collected about the students themselves or from the web. For the most part, completion of the handout was student driven, though I supplied substantial structure on how that discussion would proceed. Students were encouraged to bring a calculator each day, but I did not instruct them on any particular calculator type. While examples were interspersed throughout the class periods, I would devote about every third or fourth day exclusively to examples that students worked on in groups. They would move their chairs together and I would walk around answering individual questions, providing a summary at the end of the period.

Students submitted weekly homework assignments, due on Friday. They were encouraged to work together but needed to submit individual write-ups to the problems. I selected 8–10 problems from the text and from other resources (e.g., *Workshop Statistics*, Rossman and Chance, 2001). Solutions to the problems would be posted on the internet over the weekend and then each Monday students would take a quiz concerning this previous assignment. The primary motivation for these quizzes was to give students examples of the types of questions they might see on exams, as these problems were often more conceptual and required more explanation than what they had done on homework assignments.

A part of each homework assignment (3–4 problems) was labeled “practice problems,” and on Wednesdays students were chosen at random to present solutions to these problems at the front of the class. Throughout the semester, each student was expected to do one such presentation. The presentations were not graded and I emphasized to students that the solution did not even need to be correct and that the presentations were to be used to begin classroom discussion. The primary goal here was to give students practice presenting their work, to provide some motivation to start the assignments early in the week, and to encourage more classroom discussion of topics most relevant to the upcoming homework assignment.

So far I have left out one crucial detail. After I arrived at UOP, we reintroduced a weekly one hour lab period. We had a computer lab with 24 machines and students were given several time slots on Thursdays in which they could enroll (so each class section was split roughly in half). Technically this meant students were in class 5 hours a week and that the instructor spent 6 hours per week for each section of the course, though some faculty then reduced the length of the regular lecture time period. During this lab hour, students predominantly worked in pairs on the computer, there was very little to no formal instruction time. The lab instructions guided students through the computer commands (predominantly Minitab) and asked them to make notes of different observations and output, and then they were to respond to a series of questions designed to help them focus on the conceptual ideas behind their observations and tie the material together. The goal was for students to complete the computer-intensive portion of the lab during the one-hour period, and then complete the lab write-up with their partner outside of class, due the next lab session. A binder containing the lab assignments for the semester was distributed the first day of class (with funds supplied by a new lab fee that students paid when they registered for the course). The initial labs were very directed, with students answering specific questions, but as the semester progressed, the instructions of subsequent labs become increasingly open-ended and students were expected to work through the material more independently, building on what they learned in earlier labs.

For about half of the 14 labs (for list of labs, see www.rossmanchance.com/uop/chance/lab_notes.html), students were expected to complete an experiment or observational study to collect the data they would analyze (e.g., drops of water on pennies, measurements of heart rates, Coke vs. Pepsi taste test, comparison shopping at two local grocery stores, breakdown of M&M colors, gummy bear launching). Often the data collection was done during class time (e.g., all but comparison shopping in the above list). The other labs explored concepts using genuine data (e.g., fan cost index, attendance at major league baseball games, finger lengths of criminals) or used conceptual simulation software (e.g., www.gen.umn.edu/research/stat_tools/). When students collected the data themselves, they completed a longer lab write-up, organized as a technical report that including the following sections: Introduction, Data Collection Methods, Results, Discussion, and Conclusion (see Appendix).

What of the above is innovative? I feel students were asked to do a fair amount of writing and explanation of their solutions and that the content and pedagogy was in line with modern recommendations (e.g., ASA/MAA Joint Committee Recommendations, Cobb (1992)). Students used technology as a tool and were never asked to memorize formulas. They soon learned they had to justify their answers and a final numerical result was seldom sufficient. Students always worked with genuine data and I attempted to invoke a real-life flavor as much as possible. Often they were asked to analyze data from current news articles as well. Lectures were highly structured but activity driven, and more Socratic in nature.

And now

After moving to Cal Poly and the quarter system, I learned time is an even more precious commodity and I have not yet found the optimal balance of various class activities. The student audiences, course prerequisites, and content coverage are quite similar but we meet for many fewer days. The cost has been a reduction in student presentations, written assignments, and "fun" activities (including guest speakers and videos). I am hoping to change this as I feel these components are essential for developing the deeper understanding that I strive for.

The primary benefit of my new course structure is that I am meeting each class session in a computer laboratory. For 48 students we have 24 machines. This allows more immediate incorporation of computer explorations and analyses directly into every class meeting. I also spend even less time on hand calculations before moving to the computer. To present material to the students I typically use PowerPoint slides, which students have access to outside of class, though we also have an opaque projection system. The seating arrangement allows for students to easily work in groups of 2 and 4, enhancing possibilities for collaboration and group discussion. I have also expanded my use of the web, both for interaction with students and for course management. Minitab is still the primary software of choice, but there are, as always, some complaints about its lack of accessibility outside of class.

In this environment, I initially shifted to a more constructivist, activity-oriented structure. Students spent much more time working through activities, with my lecturing coming predominantly at the beginning and end of the class period. I have tried this in 50 minute classes, 70 minute classes, and 2 hour classes. I am currently shifting back to a bit more structure during the class, especially for the longer class meetings, as I find student's attention span for either lecture or activities is about 20 minutes. The main components from UOP that have stayed consistent are the use of student writing, periodic quizzes, and term projects. I have found success with more use of "paired quizzes" with two students turning in one set of answers. This reduces my grading load as well as adding more student discussion and debate to their experience.

A typical class

The following describes the 110-minute introduction to numerical summaries that I used last year at Cal Poly. This was the fifth lecture in the course, following 8 hours of discussion of data collection methods and graphical methods for displaying data. I began by posting a cartoon dealing with a character's use of the term "typical" that I scanned in from my local paper and then I spent a few minutes reviewing graphical summaries. I recapped what I wanted them to most remember from a *Workshop Statistics* (Rossman and Chance, 2001) activity they had completed at the end of the previous lecture (Activity 3-5 concerning choice of interval width in histograms), answered questions about using Minitab, and demonstrated a java applet that illustrated the main points of that activity.

This lesson began with standard definitions of mean, median, and mode, including notation. On their handouts students had the 12 average monthly temperatures for San Luis Obispo that I downloaded from the web and we stepped through how to calculate these numerical summaries. They then practiced using some very skewed data as I and a teaching assistant walked around the room answering individual questions. We then compared our results as a class and discussed why the mean and median behaved so differently in this case, the effect of outliers on each measure, and how the mode is not always very useful for quantitative data.

Next, we discussed the need for measures of spread. I tell the students that I had considered jobs in San Francisco and Washington DC when I left grad school and wanted to judge the cities based on their temperatures. I ask them if the fact that these two cities have similar average temperatures makes them equivalent places to live and students begin to tell me the need to consider the variability in the temperatures. We define the range and they tell me its nonresistant downside. Then we define the standard deviation. My goal is to show students enough of how the standard deviation is calculated for them to understand what it measures, but I do not have them practice this calculation by hand. Instead I focus on its properties (e.g., nonnegative, measurement units) and disadvantages. Then students work through Activity 5-1 in *Workshop Statistics* to calculate the interquartile range, five number summary, and boxplots. They also begin Activity 5-7 which has them match standard deviation values to different histograms and address prevalent misconceptions of variability such as "variety" or "number of distinct values."

In the remaining time we discuss common misconceptions from the first quiz that was just handed back, students are reminded to turn in their first project report that day and they begin working on their second lab which has them using Minitab to do numerical and graphical summaries of Major League Baseball's most recent Fan Cost Index numbers. The next class period will begin by reviewing the definitions, discussing some "practice problems" from the weekly homework, reviewing the lessons from Activity 5-7, and highlighting some common mistakes I anticipate students could make, before we proceed to the normal distribution and more time on their lab assignment. Ideally there would be more time to allow students to make some of these mistakes, as they often do not remember what I only tell them.

Student assessment

In addition to weekly homework assignments and quizzes, students are given two midterm exams and one cumulative final exam. Even on these exams I strive to place more emphasis on conceptual understanding and explanation than on calculations. Exams are closed book but students are supplied with the necessary formulas and tables. Thus, students need to know which is the right procedure to use in a particular situation and how to interpret the output of that procedure, but are not required to memorize the calculation details.

At UOP, I often gave a two-part final, with one portion being a set of take-home questions. Each student was individually asked to answer three questions about a different data set. Questions were phrased as "is there an association between…" or "is there a difference in the mean …" and students were expected to conduct the exploratory data analysis, diagnostic checks, and statistical inference procedures required. Each student has a different data set and they were expected to determine the appropriate procedures based solely on the phrasing of the research question and on the nature of the data being analyzed. Students were given a week to use their notes and Minitab to complete the analysis. They could also "buy hints" from me (though Minitab hints were typically free). I told them I would charge them points if they felt they wouldn't be able to do later parts. Most students accepted the challenge rather than lose the points.

In my mind the most important assessment tool is the group project (see e.g., Fillebrown, 1994; Macisack, 1994; Ledolter, 1995). In fact, I consider this as much a learning activity as an assessment, and it is the one part of the course I vowed not to relinquish in my switch to the quarter system. Students are asked to work in groups of three to five to develop their own research question, design a data collection strategy, collect and analyze the data, produce a written report, and give an oral presentation to the rest of the class. I require periodic reports throughout the term, but otherwise students are responsible for proceeding through the project essentially on their own. Through these periodic reports I provide feedback and ask questions to help them refine the study at each stage. The second project report (their data collection plan) is peer reviewed by other students in the course as well and the feedback provided before they begin collecting any data. The oral presentations are 10–15 minutes during the final week of classes and many students now use PowerPoint presentations. These are also peer reviewed and students are given this feedback as part of their final project evaluation.

I strongly believe that these projects are the most valuable learning experience for my introductory students. These projects engage students in the entire process of statistics, often expose students to the difficulties encountered in collecting data and to the perils of messy data, and give them an avenue for demonstrating their abilities to use the tools they have learned in the course and to decide which tools are the most relevant for answering a question of their choosing. Students also often comment on how much more they have learned in having to apply their knowledge to a question that interests them.

In designing assessments to use in my course, I try to make the assessment process iterative—giving students feedback they can utilize on subsequent assignments in the course. For example, for lab write-ups I often try to minimize the level of detailed instructions I provide, preferring students develop their own style instead of following a template. However, I do post a high-scoring student report for other students to review before their next lab assignment. This provides students with feedback (based on the work of their peers) during the learning process, instead of only at the end of the course when it is too late for them to make any adjustments, while also encouraging their creativity and individuality. Then they apply the lessons from these earlier labs to later assignments and to their final project report which is essentially a slightly more extensive lab write-up. Similarly, the quizzes provide feedback on the type of questions I will be asking, and the types of answers I will be looking for, on their upcoming exams.

In grading student work, my philosophy is to be encouraging, given the need to reassure and motivate statistics-phobic students. I give substantial partial credit if they attempt a problem, as getting them to take that first step is often the hard part. Especially with the projects I feel the experience they gain in doing the project should be a factor in their final grade. Still, especially in the lab write-ups (which I grade myself), students are graded on their ability to communicate and to support their conclusions based on the empirical evidence. For example, with the full lab write-ups at UOP, 20 points were allocated for completing the output, 25 points for their discussion and 5 points for the quality of their written report. The lab assignments at Cal Poly can be found at statweb.calpoly.edu/chance/oldstat217/labs.html, which includes a few model labs.

One thing I quickly learned in using more class activities is that students need to be made aware rather early in the course that they will be responsible for understanding the concepts covered in the activity. Too often students see these activities as merely "play time" and do not sufficiently reflect on the ideas they are learning. By requiring students to turn in summaries of these activities and by writing exam questions that are directly related to them, I hope I have conveyed this message very clearly.

An early activity in the course is "matching histograms to data." I adapted an *Activity-Based Statistics* activity that gives students a set of histograms and a set of variables and asks the students to choose which variable corresponds to which histogram by focusing on features such as shape, granularity, and perceived spread. My modification is to use data that students supplied in an initial course survey on questions such as their gender, height, soda preference, the amount they spent on their last haircut, number of siblings, and guesses of my age. Students find this quite engaging and can often look around the room for more information, further personalizing the experience.

Students turn in their explanation of which variable they believe goes with which graph. Then on a 5 point scale I evaluate how well they supported their arguments. I am less concerned that they make the exact matches but that they support their answer (especially for the pairs of graphs with similar shape), that they correctly use the terminology we have been discussing, and that they don't ignore some essential features. For example, gender and soda preference are the only graphs with just two categories, but how do they justify which has a more inequitable split? Also, do they think about the zero values in discussing cost of haircut vs. number of siblings? Students generally receive high scores on this early activity but they also learn that they will be responsible for the concepts discussed in the activities. I often follow this up with a similar matching activity for different variables on their first exam.

Assessment of the course

I am afraid I have been guilty of relying primarily on student evaluations and anecdotal feedback to help me gauge the effectiveness of the courses I teach. However, I do attempt to adjust my course substantially based on student feedback. This feedback ranges from details about the instructions of individual assignments to how I structure time spent in and out of class. I believe that I construct fairly challenging exams so that any student who does well on my final exam has developed significant understanding of the concepts. I am often impressed, as are the students, at the quality of their final projects. These projects clearly reveal what students are now able to do with the statistics they have learned. At UOP, when I had time to show the last *Against All Odds* case study (COMAP, 1989), students would often remark on how much of the information in the video they now understood and that it really reinforced the amount of material they had learned during the term.

There is a delicate balance between challenging the students enough and overwhelming them, especially in a course that students are not taking voluntarily. I have conducted some informal analyses of student reactions to the course and student retention. I found that, as I refined some of the assignments and achieved better balance at UOP, student appreciation of the course components as learning tools did increase. Examining student ratings over time has shown a steady improvement in students' opinions of different course components. For example, the graph below shows a summary of students' open-ended comments about the lab component of the course over six semesters. The percentages of students who commented that the labs were "helpful" or "time consuming" or who gave a predominantly negative response are shown in Figure 1.

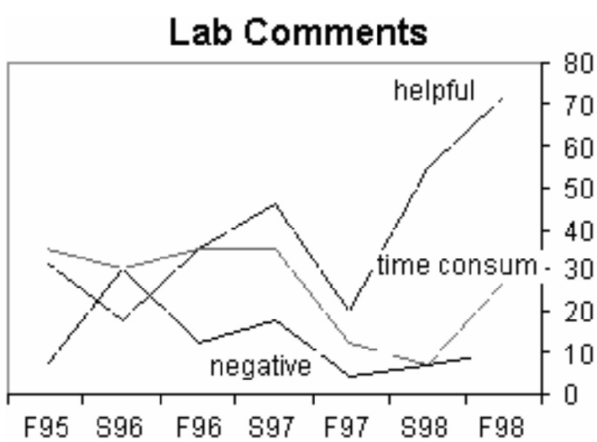


Figure 1. Student feedback on lab component of course

After I had taught the course for 5 years, a majority of students were now saying that the homework assignments, lab write-ups, and projects were beneficial in their learning. They still thought the workload was heavy, but perhaps my best reward at UOP was when students would return a year or two later, after seeing statistics again in their upper division courses, and comment that they felt more prepared than their peers who had been in other statistics courses. Students often also return and tell me how they are no longer able to view advertisements or news articles without being constantly skeptical about how the data were collected. I also felt I was doing a good job when a common student "complaint" was that I was "making them think" too much.

To conduct assessments on a larger scale, I received an Irvine Assessment Grant at UOP. One component of this project was to send a random sample of students a follow-up survey and to invite a few to complete selected items from Garfield and Konold's SRA tool (Garfield, 1991). The survey asked several questions to assess students' attitudes about statistics and one question required them to apply statistical knowledge. Some of these responses were compared to those of current students at the end of the semester. One finding is that some misconceptions that are not as present at the end of the semester do reappear over time. For example, the misconception that "good samples have to represent a high percentage of the population" was more prevalent among former students than current students. Thus, it is paramount that faculty work together throughout a student's college career to reinforce these ideas in subsequent courses.

Other components of this assessment project included external evaluation by statistics educators of course materials and student work and observation of student time-on-task in the classroom and lab. The latter revealed that students appeared to be more engaged during in-class activities and labs, though only for a finite period. This has lead me to "mix it up" more, spending time on lecture, discussion, and activity during every class period.

I have also attempted to build communication bridges with the client disciplines and directly ask colleagues in other departments if they feel the course is adequately preparing students and meeting their needs. Such conversations are crucial to a successful, meaningful course. I have found when I ask about course goals, faculty in other departments often assume I am looking for a checklist of topics, so care is needed in how the questions are posed. Faculty in other disciplines tend to have similar goals to what I've described, but are not always convinced that these goals are being met, so it may be useful to discuss what is currently going on in the courses and how that might differ from what was done in the past.

Developing an innovative statistics course

I was very lucky when I started at UOP in that I had tremendous freedom in how I developed the course. At first it was a little too unguided, but later I appreciated the opportunity to choose my own textbook, create the extra lab hour, and develop my own lecture style. I also had a fair amount of freedom in the content I covered, though I believe that statistics courses are ultimately responsible to provide content appropriate for the client disciplines from which the students come. The only real obstacles I faced were time and money. Developing activities, integrating them into the course, making sure there is a coherent theme throughout the course, finding examples that cater to student interests, searching the web for interesting data sets with the desirable features, grading significant portions of underdeveloped student writing—all required major amounts of my time.

As a new instructor I was given a reduced course load my first year and this was critical to the course development. I also received a Planning and Development Grant to help initiate development of my course laboratory manual. It was extremely helpful to me to be able to travel to professional conferences and workshops to learn more about teaching this way and to share ideas with other statistics educators. This was especially important to me as the only faculty member at UOP whose primary interest and training was in statistics.

In 1995 I attended a meeting for "isolated statisticians." The discussions were quite valuable as were all the contacts I made there. At this meeting I began working with Allan Rossman, first on hosting one of his STATS workshops at UOP, then collaborating on curricular ideas. At the Activity-Based Statistics workshop I met Bob delMas and began using his computer software and sharing feedback. At the Chance workshop I met Joan Garfield and was able to personally request copies of hard-to-obtain papers which she provided. I have been very fortunate to have had these opportunities to reflect on my teaching with experienced education reformers and that some of these initial contacts have developed into long-term collaborations.

The positive response to my innovations from students was immediate. At first I was concerned that I would not have the funding to purchase the materials needed in activities such as counting candies or gummy bear launching. However, the overall expense of buying candy and other food for activities was less than I would have predicted and seemed to be worth it, given the favorable student responses to these activities. It can also be costly to the instructor or the student to incorporate supplementary materials (such as lab manuals and daily lecture handouts), but I feel the additional gain in student ability to use the materials as reference and the time saved in class from copying examples from my paper to the board to their paper is immeasurable.

It is important to realize that this style of teaching can be quite demanding, not only in time and money, but also in terms of my own energy, and despite all the time and effort involved, not all student response will be positive. In fact, I have observed a trend that suggests my student evaluations are not particularly high the first time I teach any course. However, as I am able to better integrate materials and better convey my expectations to students up front (and get more sleep), the evaluations gradually improve. Thus, it is important for faculty to have the support of their department and administration when trying to develop more innovative teaching. Such developments may initially lead to negative returns, but it is the long term returns that they need to focus on. Departments should also consider the additional amount of grading time needed for conceptual questions and written reports.

Future plans

My courses at Cal Poly are still under development as I modify them to fit my teaching beliefs and still meet department requirements. I hope to obtain better integration of the required course content with activities to enhance learning, such as student presentations, video clips, and guest speakers. In particular, I hope to improve my ability to motivate the topics for the students. I often survey students on their interests and on applications they may see after my course, but I am still in the process of finding ways to integrate this information into the course. I hope that as the structure of the course becomes more stable, I can spend more time scanning recent articles from the news and from other disciplines for statistics examples and applications to motivate students. I am also hopeful that as students enter my course with higher levels of computer literacy, I will be able to incorporate more technology, both in class and outside of class. However, this may require use of a more portable platform (e.g., more java applets on the web) or a change in computer server policy at my campus so that students will have more access to the statistical software and conceptual demonstrations.

References

Cobb, G. (1992). Teaching statistics. In L. Steen, (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 3–43). Washington, DC: Mathematical Association of America.

COMAP— Consortium for Mathematics and its Applications (1989). *Against All Odds: Inside Statistics*, Annenberg/CPB Project, www.learner.org/resources/resource.html?uid=65&sj=MATH

Egge, E., Foley, S., Haskins, L., and Johnson, R. (1995). *Statistics Lab Manual* (3rd ed.). Department of Mathematics and COmputer Science, Carleton College.

Fillebrown, S. (1994). Using projects in an elementary statistics course for non-science majors. *Journal of Statistics Education*, 2(2). Retrieved October 9, 2003, from www.amstat.org/publications/jse/v2n2/fillebrown.html

Garfield, J. (1991). Evaluating students' understanding of statistics: Development of the statistical reasoning assessment. In R. Underhill (Ed.), *Proceedings of the Thirteenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Blacksburg, Virginia, 2, 1–7.

Halvorsen, K.T., & Moore, T.L. (1991). Motivating, monitoring, and evaluating student projects. In *American Statistical Association Proceedings of the Section on Statistical Education*, Atlanta, GA, 20–25.

Ledolter, J. (1995). Projects in introductory statistics courses. *The American Statistician*. 49(4), 364–367.

Macisack, M. (1994). What is the use of experiments conducted by statistics students? *Journal of Statistics Education*. 2(1). Retrived October 9, 2003, from www.amstat.org/publications/jse/v2n1/mackisack.html

Moore, D. S. (2001). Statistics: Concepts and controversies (5th ed.). New York: W. H. Freeman.

Moore, D. S., & McCabe, G. P. (1999). Introduction to the practice of statistics (3rd ed.). New York: W. H. Freeman.

Rossman, A. (1996). Workshop statistics. New York: Springer-Verlag.

Rossman, A., & Chance, B. (2001). *Workshop statistics* (2nd ed.).Emeryville, CA: Key College Publishing.

Scheaffer, R., Gnanadesikan, M., Watkins, A., & Witmer, J. (1996). *Activity-based statistics*. New York: Springer Verlag and Emeryville, CA: Key College Publishing

Spurrier, J. D., Edwards, D., & Thombs, L. A. (1995). *Elementary statistics laboratory manual*, DOS Version, Belmont, CA: Duxbury.

APPENDIX

Guidelines for full lab reports

Some of the labs require "full lab reports." The goal is to learn how to effectively communicate the information you have acquired. In any discipline, all investigations require clear and efficient communication of the results, otherwise the knowledge you have obtained is useless! However, communication skills are consistently rated as one of the three main weaknesses industry executives see in recent college graduates, (along with teamwork and statistical skills). The main features of good reports are clarity, logical organization, succinctness, and clear labeling. The following should give you some initial guidelines for writing these reports.

For these labs, your write-up should include the following sections: I. Introduction, II. Data Collection Methods, III. Results, IV. Discussion, V. Conclusion, and possibly an Appendix:

Introduction: The Introduction should be worded carefully, as that may be as far as some people get. The goal is to convey to the reader the goals and most essential outcomes of the study. It's often easier to write this section *after* you've written the rest of the report so you know what you are trying to summarize. This summary should include your motivation for the study and why the reader should continue to read your report.

Data Collection Methods: The goal of this section is to tell the reader precisely how the data was collected. This should be a description of the process, not the data. This description should include details of the measuring instruments used, operational definitions of basic measurements, randomization, precautions etc. There should be enough detail that someone could replicate your study.

Results: The goal of this section is to describe the data measured and the analyses performed. You should use tables and graphs to *summarize* the data—an effective summary is worth a thousand data points. You should also include any changes that were made to the data (scaling, transformations, missing values, points that had to be discarded). You should include a "map" telling the reader how to find the information. Tables and Figures need to be labeled and numbered for easy reference. Each figure or table should be able to stand on its own and tell its own story (units!). Often Minitab can make the annotations for you, but sometimes you may want to write in a title or label by hand or with Word. Sometimes you can use computer output as is, and sometimes you will want to select out relevant information and construct your own summaries (see examples below).

Next you should describe the analysis methods. Be complete in your description so that the reader can assess the validity of your methods. You need to include a description of what was being tested, the statistical methods used, why they were appropriate, and any assumptions (and verifications) that were made. You may even want to indicate what Minitab commands were used. *You don't need to include all your output, but select the output that supports your conclusions.*

Discussion: Now you get to interpret all the above results. You should include any explanations you may have for what you have found in the data. Recall any problems you had collecting the data (e.g., is it really a random sample?) and how your interpretations may be limited. It is also possible to combine this section with the previous one, with Minitab output interspersed through the discussion.

Conclusion: Briefly summarize your report. What is your final answer to the question? What are the implications of the results? Use non-statistical language. Include any ideas or suggestions you have for future experiments.

Appendix (Optional): If your raw data is too extensive to be included in the report, it may be placed in an appendix (again, well labeled) when necessary. Any equations or technical details can also be placed in an appendix.

The text part of your report should be about 3–4 pages in length, never more than 5. As the labs become more open-ended you will be responsible for selecting appropriate methods and justifying their use. You should be ready to review earlier labs, notes and material in the textbook.

This type of writing takes practice; remember, we don't expect perfection the first time out! I will provide you with some brief examples of a couple of good and bad techniques. Exemplary write-ups will be placed anonymously on the computer lab bulletin board and on the course web page for your review.

Section 4

Teaching Statistics to Students in Other Disciplines

13

Teaching Statistics as a Language to Students in Psychology

Karl N. Kelley
North Central College

Too often students believe that statistics is something that is "done unto them" and their goal is simply to get out of the course with a passing grade. Therefore, I try to encourage students to take ownership of their own learning and to have fun with the material.

Becoming a teacher of statistics

I had just finished my first year of graduate study in psychology at Virginia Commonwealth University when the assistantships for the following year were assigned; my assistantship consisted of teaching several laboratory sections for the behavioral statistics course. In this course, there would be one large lecture section taught by a faculty member and several laboratory sections taught by graduate assistants. My assigned role in this course was to answer student questions about the lecture and to apply the information covered in class to solving various problems.

Although I was somewhat nervous about teaching my first course, I quickly realized that the students' fears were considerably greater than my own. What I discovered in teaching this course is that the students often became so obsessed with "getting the right answer" that they failed to see the implications of the problem. They had the belief that if they could just grind out the right answer, everything would be OK.

I began to notice that although many students could calculate the correct answer, they often did not know what to do with the results. For example, they could compute that there was a statistically significant difference between two groups, but many could not explain how to use that information to make a recommendation about a treatment plan or other decision. At that point, I started developing problem sets that would demonstrate to psychology majors that the results of a research project have direct consequences to clients. Misinterpreting the results of a study can endanger a person. Statistics are not just about getting the right answer in a computation, but using that information in a reasonable manner. Thus, after completing the problem, I asked students to explain to a fictitious client why they would suggest a particular treatment.

The experience in teaching my first statistics class is congruent with my professional training and background. As a graduate student in psychology, I took the traditional univariate and multivariate statistics courses my first year. In subsequent years, I also elected to take several advanced statistics courses offered in psychology and other disciplines. There was a computational aspect in each of these courses but my training emphasized the practical use of statistics. My instructors downplayed the mathematical intricacies of the analyses and focused on applying some basic principles in probability and sampling. I began to take ownership of statistics when I began to do my own research and had to explain my decisions to others—including my choice of analysis and interpretation of the results.

In both my undergraduate and graduate statistics courses, I found I had to work very hard to understand the material. In addition to practicing many problems, creating note cards with terms on one side and definitions/examples on the other, I also joined or started study groups. Although time consuming, these strategies helped me master the material and provided me with a sense of accomplishment and competence. As Ptolemy argued in the *Almagest*, the study and practice of mathematics (and I would include statistics here) allows one to see everything in life more clearly because mathematics imposes on us consistency, order, and symmetry—ideas associated with understanding and the divine (Toomer, 1984).

Teaching statistics continues to be one of the most challenging and rewarding experiences of my academic life. The challenge in presenting this course stems from the discrepancy between my views of the course and the students' perceptions that this course is another math class they need to survive. My bias in teaching this course is my passionate belief that to be successful in the world today, individuals must be able to use and consume basic statistical information. I treat this class as the development of critical thinking skills that will help students make better decisions, which will enhance both their career and personal success. Yet, despite what I see as the intrinsic value of this subject, students are often tentative in embracing the material. As a teacher, I perceive my job as not simply conveying information, but convincing students that the material is important. If I am successful in showing them why the information is valuable, they are more likely to engage in behaviors that will lead to understanding (and consequently a good grade in the class). Thus, I seek ways to present this sometimes complex material in an understandable and interesting manner.

The rewards come from the students who discover that they actually like the topic and find that they understand more than what they anticipated. Rewards also come from occasional letters I receive from graduates who tell me that they are finding a use for what they learned in my class, or they were surprised to find that they were able to explain a statistic to a colleague or supervisor.

I currently try to teach statistics more as a language course than as a mathematics course. My focus is to treat statistics as tools that help us understand a complex world by assisting us in evaluating evidence. Although there are important mathematical principles at work, the primary objective for my course in introductory statistics is to ensure that students can explain and use the material. Thus, I engage students in statistical conversations as opposed to focusing solely on statistical computations. The dialogue created in these conversations engages students and fosters active learning.

My course

I teach statistics at North Central College, a small liberal arts college (approximately 2500 students, of which 1800 are traditional undergraduates and 500 are graduate students) located about 35 miles west of Chicago. Currently, the college offers four different statistics courses: two calculus based courses offered in the math department, a business statistics course, and behavioral statistics. The behavioral statistics course (PSY 250) emphasizes research methods and hypothesis testing more than the other introductory courses. Prerequisites for the course include college algebra along with computer experience, specifically with spreadsheets and databases.

One of the unique features about North Central College is that we are on a 10 ½ -week trimester schedule. Most courses (including statistics) are worth 3 semester hours of academic credit and meet either 3 days a week for 70 minutes each day or 2 days a week for 1 hour and 50 minutes a day. Thus, we cover the same amount of material over a shorter period. Full-time students are allowed to take either 3 or 4 classes a term, so they are typically taking fewer classes at any given time than in a traditional semester.

The target audience for the PSY 250 Statistics course is primarily social and natural science majors. It is a core requirement within psychology and serves as a prerequisite for Research Design & Experimentation, Cognitive Psychology, and Tests & Measurements courses. This course also fulfills one of the core requirements for a B.S. degree. Although the actual distribution of student majors changes over sections, typically about 50% of the students taking this course are psychology majors, 25% are natural science majors, and the remaining 25% consists of a variety of other majors including elementary education, English, philosophy, etc.

The course has a maximum enrollment of 30. In most years, the department offers at least four sections, most of which fill to capacity. Therefore, in any given year we teach about 120 students; I typically teach two sections of this course a year and other members of the department alternate teaching the other two sections.

Although there is some variation in how instructors organize the course, we have agreed on several common goals:

- To address some basic methodological issues in research design
- To have students be able to select and compute an appropriate statistical analysis
- To make appropriate inferences based on data
- To become a critical consumer of statistical information
- To be able to use a computer package (SPSS) to analyze data

Ultimately, we are training students to detect and interpret patterns of data that support or fail to support identified theories. We cover three basic techniques for seeking these patterns: correlational, experimental, and quasi-experimental designs. We stress the identification of variables (dependent, independent, and subject) and measurement (operational definition). From there, we cover basic descriptive statistics, probability, and then inferential statistics. Specifically, we teach the following tests:

- One sample *z*-tests
- One sample *t*-tests
- Two sample *F*-max tests for variance
- Two independent sample *z*-tests
- Two independent sample *t*-tests (both pooled and separate variance computations)
- Two dependent sample *t*-tests
- Correlations (Pearson and Spearman)
- Simple regression
- One-way ANOVA with post hoc (Tukey & Scheffe)
- Chi-Square (goodness of fit and independence)

Our hope is that we can convince and demonstrate to students that understanding statistics can help them become more effective decision makes and better consumers of information. If we are successful in doing this, the students will in turn be more likely to invest the resources they need to in order to understand the material.

A typical class

Apart from teaching some specific statistical techniques, I have several general goals that permeate all aspects of the course. These goals include:

- Teaching statistics as a language rather than a math class
- Fostering critical thinking and reasoning skills
- Encouraging student responsibility for their own learning
- Using technology to analyze and present data

I begin this class by describing what I refer to as "conversational statistics." My goal is to teach the use of statistics to communicate or to use statistical information to make informed decisions, so I want students to be able to use some statistical terminology in a consistent and coherent manner. For example, when someone says that there is a statistically significant difference ($p < .05$) between group A and group B, does that suggest that everyone in group A is different from everyone in group B? I want students to be able to explain variability and to understand the concept of confidence and errors.

Although I emphasize the conversational aspects of the course, there must still be numbers involved in our conversations. Upon seeing these numbers, many students then immediately assume math—and for some, that association is accompanied by negative emotions. Many students in our behavioral statistics class do not feel confident in their mathematical abilities and are not interested in learning more about math. In 2000 and 2001, I surveyed students taking statistics and found that 61% reported that they believed they had low ability in math. A slightly higher percentage of students indicated they would not take more math courses than are required for graduation.

I do not in any way wish to suggest to the students that mathematical skills are not important, but most students are not going to become statisticians. Indeed, most graduates will not need to compute statistics in their jobs or personal lives. Of those who will perform statistical analyses, most will use statistical software to obtain their results. Nonetheless, I believe it is important for students to understand at least some of the computational basics for statistics. To understand statistics, students should be able to compute means and standard deviations. Although we perform some computations in the class, if students can add, subtract, multiply, divide, understand order of operations, understand some basic algebra, and use a calculator to square a value and find a square root, they should be able to handle the math part of the class. The challenging part of the class is selecting the correct analysis and making an appropriate inference.

To foster my emphasis on language, I begin most classes with a conversation about statistics. In the beginning of the course, the conversation focuses on student impressions about statistics. As Reid and Petocz (2002) suggest, there are six different ways students can conceptualize statistics. These conceptualizations range from thinking of statistics as a set of individual numerical activities to viewing these analyses as tools helping us make sense out of the world. By getting a sense of the students' schemas, teachers can modify the course to address some of the issues associated with each conceptualization. For example, students who view statistics as a purely mathematical activity may excel at computing and analysis but miss the implications of the project. On the other hand, students who focus more on the application of statistics may struggle with the actual computations but will be able to effectively use the information to make decisions based on the data.

As the course proceeds, I present students with some basic terms and short problems, asking them to explain their answer to someone who is not an expert in statistics. When defining terms, I encourage students to develop examples relevant to their own interests and areas. For example, I explain the difference between independent and dependent variables in class using an example. When I ask students to explain these terms later, there is a tendency for them simply to repeat my example. Although this demonstrates that they have learned my example, they may not really understand the concept until they develop their own examples. I also give them examples where, in some cases, the independent and dependent variables have been incorrectly identified. Their mission is to identify these errors and correct them.

Near the end of the course, these conversations become more elaborate and sophisticated. There are times when I present students with the results of a study and ask them to explain what it means. For example, I tell them a research project reported that drug X significantly lowered the blood pressure in a sample of patients with hypertension. If I have hypertension and I take this drug, will it lower my blood pressure? Students need to explain their answer to me using non-technical terminology. Near the end of the term, these conversations can occupy between 15– 20 minutes of the class.

Apart from the conversational aspects of the course, I also try to emphasize critical thinking and reasoning skills. In many ways, this course is about evaluating evidence. I ask students to consider what evidence is sufficient to support a position or give them confidence in making a decision. For example, I give students a statement such as "Psychology Majors are smarter than students majoring in ____ (pick your favorite major)." I then tell the class that a student just conducted a survey and found that psychology majors had a higher GPA than students majoring in the other area. Students have to decide if that evidence is compelling enough to have them agree with the stated position. If it is not, they must address their reservations and concerns. As the class progresses, I present students with newspaper stories and published research papers and ask them if the data support the conclusions. My goal is for students to critically evaluate the evidence before accepting the premise.

A third general theme that runs through my course is encouragement for students to take responsibility for their own learning. Although I will address this issue more fully in the assessment section, this theme is important in the everyday conduct of the class. Students have often entered this class viewing statistics as a passive process. They tend to see statistics as something done unto them rather than something that they need to do for themselves. As a teacher, I see myself not just teaching a content area, but also empowering students to learn how to learn, and I want them to gain the confidence that comes from mastering a difficult subject. Specifically, I expect students to come to class prepared. If they do not understand a problem or concept, I encourage them to take advantage of the resources available to them, which includes visiting me during my office hours for assistance. Students can also visit a web site (created by the authors of the textbook) that has an online quiz for each chapter. The college also provides tutors for students in statistics. Thus, students need to decide what it will take for them to achieve the level of success that they desire. For some students, this may be a minimal amount of time outside of class whereas for others it may be considerable.

A fourth theme that runs through the course involves the use of technology. Since I have been teaching statistics, I have increasingly incorporated more technology into both the classroom presentation of material and the graded components of the course. Specifically, I have used the internet (including a course web page and web based demonstrations and assignments), PowerPoint, Word, and SPSS.

In addition, for the past year I have maintained a course web site using the Blackboard (www.blackboard.com/) system. I have found this system to be very useful to both my students and to me because it allows us access to the course from any internet connection. On this site, students have access to basic course documents (syllabus, assignments, class handouts—all in PDF format), PowerPoint lectures (in PDF format), their grades, and a variety of Web resources (including online quizzes psychology.wadsworth.com/book/gravetterwallnau5e/quizzes/gravetterwallnau5equiz.html, general resources such as Hyperstat Online davidmlane.com/hyperstat/index.html, and even links to statistical humor www.ilstu.edu/~gcramsey/Gallery.html). Blackboard also provides students with a way of communicating with me and other students in the class via e-mail. Another interesting feature of this product is that it allows the instructor to monitor student use of the various content areas. For example, based on these reports, I can easily tell that almost every student checks their grade reports several times a term. Most students use the e-mail feature but very few use the general resources or statistical humor sections.

In the class, I use several web-based demonstrations (e.g., guess the correlation www.stat.uiuc.edu/~stat100/java/GCApplet/GCAppletFrame.html and hypothesis testing acad.cgu.edu/wise/hypothesis/hypoth_applet.html). These fun and interactive tools illustrate complex concepts. Students can readily see the difference between a correlation of .93 and .66 using similar sample sizes. After covering these in class, I suggest students try the projects at home. Although these are not required projects, I believe that they can be very beneficial to students. However (using data from Blackboard that monitors student use of the internet) only about 40% of my students spend over an hour on these tasks—time well below what I think would be beneficial for many of them. However, I have chosen not to require these assignments because some students do not need to do them in order to understand the material whereas others do. This issue relates to my theme of taking responsibility for one's own learning. Students need to decide if spending extra time with these assignments is beneficial or if they need to focus their attention elsewhere.

During the lecture part of the class, I use PowerPoint presentations to highlight key areas of the content and to present several complete examples. I post these presentations on the course website (Blackboard) so that students have access to all of the topics and formulas presented in class. Many students will print the PowerPoint presentations before class and take notes on their copies.

Students are required to learn and use SPSS to analyze the data for their class projects and papers, described in the following section on assessment. SPSS provides the students with tools to help make better decisions about data, but the challenge for the students is in selecting the appropriate analysis and providing a reasonable interpretation of the results. Students are also required to learn how to create graphs in SPSS, save them, and then imbed the graphs into Word documents. I have found that SPSS is relatively easy to teach by using a set of structured problems, gradually asking students to use more features of the system.

Overall, technology has played an increasingly important role in how I teach statistics. Although students need to understand how statistics work, they also need to be able to use the information to make better-informed decisions. I believe that by using web-based demonstrations, students can "see" the mathematics in statistics much better. By using SPSS, students can focus more on interpreting results of larger datasets.

In summary, a typical class begins with a short statistical conversation either reviewing some basic principles or evaluating a study. There is a lecture component, which includes PowerPoint presentations and web-based demonstrations. Throughout the term, students will often work in informal small groups on problems—assisting each other with solving the problems and interpreting the results.

Assessment

I evaluate student performance using three different components: Exams, papers, and a graded study guide. Typically exams comprise about 2/3 of the final grade and papers and study guide are worth about 1/3 of the final grade.

My exams consist of three distinct sections: definition/short answer, computation, and interpretation. Since I try to emphasize the "language" of statistics, the definition/short answer section is usually worth about a third of the exam score. In this section, students define and give examples of some key terminology and respond to several short answer questions. Below are examples of the short answer items:

- What are three limitations or concerns about hypothesis testing?
- Compare and contrast the following distributions: t, z, F, and χ^2.
- Explain why we can never "prove" something using statistics.
- Under what conditions can the sum of squared deviations be negative?

Following this section, students are presented with several computation problems. Apart from computing the correct answer, they are required to explain in words what their computations mean. For example, after computing the mean, median and standard deviation for a data set, they must be able to explain the shape of the distribution and the implications for the group being studied. Although students do some calculations in this section, most of the summary information has already been computed for them. They are not expected to memorize equations and are allowed to bring formula sheets with them to the exams. As the course progresses, this section becomes smaller and worth fewer points.

The third section is the interpretation section. In this section, students are presented with processed data and are then asked to either make a decision based on that data or critique a conclusion someone else has made. As the course progresses, this section becomes larger and worth more points.

In addition to the traditional exams, students are required to write papers (see Appendix 1: Sample Project). In these papers, students are given a problem, a dataset and an audience. Their goal is to prepare a professional paper to present to this audience explaining their results and suggesting an appropriate course of action. The format for these papers requires them to have a cover page, introduction, methods, results, and discussion sections. Students are told that the introduction and discussion sections are to be directed at the target audience, whereas the methods and results sections are for professional statisticians/researchers. By having these requirements, the methods and results sections can contain the jargon of the discipline and students must use those terms and formats correctly. For example, they must be able to present an APA style t-test in their papers. The introduction and discussion sections cannot contain jargon unless those terms are clearly explained. Students are required to include the statistics they used in the report, but they must explain them to a non-technical audience.

The final graded project is a study guide for the cumulative final exam. Throughout the term, students are encouraged to work on this project as a way of studying and organizing the material in this class. The overall goals for this study guide are:

- To help organize studying for the cumulative final
- For students to see relationships among formulas
- For students to have a resource to take with them to other classes
- To have some fun

Study guides are required to have at least five sections:

- Section I: Basic Terms
- Section II: Error Terms
- Section III: Distributions & Tables
- Section IV: Formulas
- Section V: Decision Guide

Although this study guide project is graded, it is only worth 4% of the final grade for the course. By keeping the percentage for this project low, I am hoping that students are motivated to see the value of the project beyond the grades. Grades on these study guide projects are based on:

- Completeness
- Accuracy of formulas
- Correct definitions
- Creativity

In presenting the study guide project to the class, I try to stress the creativity and fun of it and in the past, some students have developed very interesting projects. Some have even incorporated cartoons, jokes, multicolor formats, and popular themes (including Sponge Bob Squarepants and Rugrats). At times, I am very impressed with the effort and quality of these documents. Although I do not have any hard data to support the usefulness of these projects, every term I have students who report that their study guides helped them in other classes. I have even had other instructors comment on the quality and usefulness of some of these projects.

The most vexing problem I have faced in teaching statistics is getting students to keep up with the material and to study on a regular basis. At one time in my teaching career, I had daily (and later weekly) graded homework assignments. My goal was to motivate students to keep up with the course and to demonstrate the value of regular study habits. For the most part, students did the assignments yet complained vociferously about it (suggesting it was busy work and took up too much time). However, the most compelling complaints came from the best students who argued that their time could be better spent in moving ahead rather than in practicing problems they already understood. In keeping my goal of having students take responsibility for their own learning, I decided to stop collecting homework assignments although I still assign daily problems, (mostly from the text book and weekly sample problem handouts). Answers for all of these problems are kept on reserve in the library, in the student tutoring room, and in my office. Students are encouraged to check their answers and if they understand the material they can move on to new problems. If, however, they struggle to understand a concept, there are additional problems for them.

I also require students to purchase a "Homework Notebook" in which to keep an organized record of all problems they attempt. They are to bring this notebook with them when they visit me during my office hours or when they see a tutor. I also ask them to bring this notebook to class every day because I sometimes select one of the assigned problems to cover in class. I can often get a good sense of how much effort students put into this course by the number of problems they attempt and where they are making mistakes (if any).

In summary, I keep the evaluation of student performance in line with my goals for teaching the class. First, for exams and papers, students must demonstrate the ability to communicate about statistics (keeping with my goal of statistics as a language). Second, to foster critical thinking and reasoning skills students are asked to resolve problems or evaluate how others have used statistics to make decisions. Third, although I regularly assign practice problems, I do not collect or grade homework but encourage students to monitor and take responsibility for their own learning. Finally, students must use SPSS and Word to analyze and present data.

Developing an innovative course

As I prepared my course syllabus for the current term, I reviewed some syllabi from years past and found that my course has changed dramatically since I began teaching in 1986. The major change has been an increase in stressing the communication of information rather than the computational aspects of the course. There has also been a dramatic increase in my use of technology for presenting information and in having students use statistical packages to analyze problems.

Apart from those dramatic trends, I try something a bit different every term. Some of my ventures work and others do not. For example, one year I required students to give oral presentations in addition to their written projects. For these presentations, they prepared a PowerPoint presentation geared toward a specific audience. I found that there were many benefits to these assignments, one being that I could ask students questions during their presentation in order to understand the logic they used to solve the problem. I believe that students learned to discuss statistical information and saw the value of making a clear and concise presentation. However, there were two significant drawbacks. First, many students were not familiar with using PowerPoint and it became "something else" they had to learn (in addition to the statistics, SPSS, and Word). More importantly, it simply took too much time. In our 10 ½ week term, there was not enough time to cover the required material and have students do 2 or 3 presentations. When I had them do one presentation, I had enough time in the course, but students did not have enough practice to prepare an effective talk.

I have been fortunate that North Central College has supported my development of this course. Specifically, I have been supported in purchasing equipment and software to use in this class. Several years ago I used a computer on a cart to take to my classroom everyday to use for demonstrations. Later, computers were installed in all classrooms and wired to the college network and the internet. Currently I teach in one of the college's "smart rooms" with a networked computer, VHS, DVD, smart board, and Elmo projector. I also teach next to a computer lab where I can assist students in learning how to use the course software, have them work on the web demonstrations, or meet in small groups to solve class assignments.

For me, the key to innovation is my own passion for teaching this subject. I continually seek new ways to engage students in this material and convey the importance of statistics. I believe my willingness to try to help the students comes through in my classes and even when ideas do not work out, my students usually accept it in the spirit it was intended.

Future plans

As I look to the future of this course, I want to find more ways to get students involved with the material. I am also seeking other ways to encourage students to allocate time in the beginning of the course before they fall behind. Although I believe I have addressed some of the issues I face in teaching statistics, there are still many I have yet to resolve. These issues include:

- Teaching to the wide variety of students in the classroom
  - Some students enter the class confident in their mathematical skills whereas others fear and dread the topic.
  - Some students are motivated to learn the material as they see it directly relating to their future careers whereas others are simply trying to finish this requirement.
- Encouraging students to keep up with their work
- Diffusing some of the negative preconceptions about statistics
- Fostering more active learning

Currently the issue that occupies most of my thinking is addressing diversity in the classroom. More than any other course I teach, I find this to be an issue in statistics. I have also found that it is often the students who are most uncomfortable with the material who are the most reluctant to seek help before they fall behind. To address this issue, I plan to invite the statistics tutors to come to class and talk about their services. I will continue to encourage students to evaluate their own needs and seek appropriate assistance.

The limiting factor in this course has always been one of time. Since the course is a prerequisite for other classes, there are some required analyses students need to learn. The more active learning strategies I use in the classroom, the less time I have to cover information. When I have asked students to read the text, come to class, and apply that information, there is often considerable confusion.

I will continue working on ways to address these and other issues since I want to be as active in learning as my students should be.

References

Reid, A. & Petocz, P. (2002). Students' conceptions of statistics: A phenomenographic study. *Journal of Statistics Education*, 10. Retrieved October 16, 2003, from www.amstat.org/publications/jse/

Toomer G.J., (translator) (1984). *Ptolemy's Almagest.* New York: Springer-Verlag.

Appendix 1

Sample Project

Project 2: Attributional Interventions — Social Sciences

The director of a research project is trying to uncover differences in how people respond to failure feedback. Specifically, she hopes to better understand how we, as psychologists, can help students overcome failure. She decides to focus her research on college students.

After conducting a substantial literature review, she and her team finds that there are three basic intervention strategies. One approach is to do nothing. A second and more traditional intervention includes sending students on academic probation to time management seminars, study skills workshops, and mandatory tutoring sessions. A more recent development (Kelley, 1996) suggests that we should address student coping ability. This includes working with student attributions (explanations of the causes of behavior). Researchers believe that by modeling healthy attributional statements ("I can control my world" "my grades will improve if I work harder") student performance will increase. To test this idea, a group of students who are currently on academic probation are randomly assigned to one of two groups: Traditional (workshops, etc.) and Attributional. At the end of the year, their GPA's are compared. For this study, the following data are collected:

Traditional			Attributional		
2.06	2.73	2.00	2.44	1.89	2.33
2.85	2.52	1.77	2.55	2.55	2.41
1.91	2.05	1.65	2.33	2.31	2.00
0.00	2.06	2.23	2.80	2.05	1.88
1.80	2.33	0.44	2.10	2.35	2.25

Audience

A residence life staff at a small liberal arts college somewhere in the Midwest.

Goal

This group is exploring options for students on academic probation. Prepare a report to the dean of students describing your project and your findings. You should be 95% confident in your answers.

Hints

- This is a professional report—it should be organized and neat.
- In the discussion, keep your audience in mind and tell them what they need to know.
- Make sure your results section includes all the necessary information and appropriate summary tables. Hint: Be sure to include appropriate critical & obtained values, significance, standard error, etc.).
- You may insert your graph into the document or include it on an additional page.
- You should be able to get all of the information on 1 to 1½ pages.

14

Teaching Statistics to Beginning Researchers in Psychology

Gary H. McClelland
University of Colorado

My two goals in designing my course are to reduce student fear to manageable levels so that students can learn, and to provide students with skills, including but not limited to statistics tools, for conducting empirical research in psychology.

Becoming a teacher of statistics

My sophomore year in high school, for reasons I no longer remember, I became very interested in experimental psychology. For a time I raised a small colony of white lab mice for conducting learning experiments in my basement, but that ended at my mother's request after one too many escapes. So as to not need a colony in the basement, I switched to human research and completed a questionnaire-based study comparing relative strengths of various psychological needs and motivations. I entered this project in the local science fair. Unbeknownst to me, there was a new award that year for the best social science project. Amidst the usual collection of cloud chambers, plant-growth experiments, and tests of detergent efficacies, there was no competition within the social science category so I won a monetary prize sufficient to purchase about a half dozen scientific books in my field of choice.

Having no idea how to select suitable books I consulted with the psychologists in the basic research group at the Menninger Clinic, which had sponsored the prize. They told me gently but unequivocally that although I had demonstrated some good scientific intuitions, my project would have greatly benefited from just a little knowledge of statistics. So, in the bundle of books they helped me select was Allen Edward's textbook *Experimental Design in Psychological Research*. Having struggled with my own data about how to decide whether a difference between groups was "real," it was a wonderful revelation that there were principled methods for making such decisions. I was hooked.

I worked my way through as much of Edward's book as I could understand (repeated measures remained a mystery), got up early to watch Frederick Mosteller's *Continental Classroom* probability lectures on television, and computed correlations, *t*-tests, and one-way ANOVAs for my subsequent science fair projects using the most rudimentary computational tools (an abacus, an adding machine, and a slide rule). I arrived as a freshman at the University of Kansas just as they were making a computer available to anyone willing to keypunch decks of cards and wait 24–48 hours to retrieve their outputs.

Desperate to avoid the arithmetic drudgery of even simple statistical tests and eager to do multiple regressions that were far beyond my computational tools, I quickly acquired basic programming skills and learned to use some early statistical packages such as Sokal and Rolf's NT-SYS. I soon found ready employment in the departments of psychology and anthropology doing statistical analyses. My cookbook knowledge of statistics exceeded what I would be taught in the required statistical methods course in my major department of psychology. I convinced them to let me substitute a mathematical statistics course from the math department. We used Hogg & Craig's textbook and I still remember the thrill when my cookbook understanding of Student's *t*-test was replaced by a thorough understanding of its theoretical underpinnings.

I went to the University of Michigan to study mathematical psychology with the support of a pre-doctoral traineeship from the National Institute of Mental Health, which required me to get an MA in either mathematics or statistics. I chose statistics and did not realize until many years later that I was one of the first graduates of this new department. After earning my PhD in psychology I almost accepted an offer from the basic research lab of a major company, but my mentor, Clyde Coombs, ardently believed, quite correctly, that I was someone who needed to teach and who would thrive working with students. He told me that "students were pure gold" and then with a twinkle added that "gold was so valuable because it was malleable."

So, for a lot less money, I arrived at the University of Colorado as an Assistant Professor of Psychology to teach introductory statistics, survey research, and attitude measurement and scaling. Believing my peers were teaching cookbook-oriented rather than theory-based statistics courses to psychology students, only because they themselves did not understand the theory, I thought I could teach a successful theory-based course using my freshly acquired knowledge. My first course was an unmitigated disaster for both me and my students. Since then I've sought to understand what statistical knowledge psychology students really need and to understand the issues—often psychological issues such as fear—that make learning statistics difficult for our students. The course described in this chapter has evolved over the subsequent 25 years as my understanding about those needs has improved.

In this era when state legislatures and governing boards are increasingly skeptical about the research mission of universities, I should note that my teaching and scientific research are closely linked. I study judgment and decision making with special attention to the difficulties normal people have with probability concepts and with making risky decisions. For example, I've published studies about couples' beliefs about dependencies in the sequence of genders born to families and how such beliefs influence decisions to have more children when their current family composition does not have the desired gender balance.

I've also studied reactions to high-consequence, low-probability risks such as those associated with living near Superfund sites (uncontrolled or abandoned places where hazardous waste is located). For a government agency, my graduate students and I have developed flood warning systems that optimally trade off the costs of false alarms and misses. My recent research pertains to the development of optimal web-based tools for aiding decisions. Examples from my research are used frequently in my teaching and of course the subject matter of my teaching is relevant to my research. Each informs the other. Hence, teaching statistics is inextricably a part of my professional career.

My course

Psychology is always among the most popular majors at most colleges and universities. And virtually all psychology departments require their majors to take an introductory course in statistics and/or research methods.

Instructors of psychology courses following the introductory statistics course expect their students will be able to comprehend in journal articles and to produce in lab reports statements such as "the mean of the treatment group significantly exceeds the mean of the control group ($t(43) = 2.35$, $p < .05$)" or "when controlling for these variables, the regression slope relating this predictor variable to the criterion variable is significantly greater than zero ($t(52) = 2.67$, $p < .05$)." And woe to those instructors of introductory statistics who do not prepare their students to be able to understand and to produce such statements.

Experimental psychologists often teach the introductory statistics courses. Thus, the typical statistics course focuses on those procedures most useful for the analysis of small-sample experimental data, specifically t-tests, one-way analysis of variance with post-hoc comparisons of means, factorial ANOVA with interpretation of interactions, and repeated-measures ANOVA. Although correlation and regression are now equally if not more important than those techniques in the field of psychology, they are not as important in the psychology lab courses and so receive some but not as extensive coverage as ANOVA designs. However, this is beginning to change and some introductory courses now even include some material on multiple regression.

Research psychologists incorrectly believe that non-parametric statistics solve more problems with messy data than they really do (e.g., most psychologists are not aware that most nonparametric tests only relax the normality assumption but still require homogeneity of variance), so typical courses cover Mann-Whitney/Wilcoxon, Kruskal-Wallis, Spearman, etc. Although probability concepts are often integral to many psychological theories and are often important for interpreting data (e.g., adjustments for guessing), probability usually receives only cursory attention.

At the same time as psychology is increasingly becoming a more hard-core scientific discipline—some departments are now almost another biology department—many psychology undergraduate students are still attracted by its "softer" side oriented to therapy, counseling, and development. These students seldom have taken any calculus courses and often are not comfortable with basic algebra. They do, however, arrive with increasingly sophisticated computer skills. Nevertheless, most psychology students dread the statistics course they are required to take.

My two goals in designing my course, given the context described above, are to reduce student fear to manageable levels so that students can learn, and to provide students with skills, including but not limited to statistics tools, for conducting empirical research in psychology. The syllabus and all other materials for the course are available on the web at psych.colorado.edu/~mcclella/psych3101h/.

The course I teach is roughly divided into two halves, with the first half devoted to learning statistical tools and the second half focused on applying those tools to answer research questions. A course week consists of three hours of lecture and two hours of lab. I have the luxury of teaching small sections of 15 to 18 honors students in a classroom in which all students are seated at individual computer stations, with a projector attached to the computer I use. However, I have previously successfully used the same basic course structure in a section of 100 regular undergraduate students with lecture in a standard classroom (with computer projection) and recitation/lab sections of approximately 20 students in computer labs. One of my colleagues has adopted my course structure and uses it for a regular 100-student section.

Even in an honors section, many students admit they are fearful. On the first day of class I address the fear issue directly and explain how I have designed all aspects of the course to reduce fear. The following are the key aspects of my fear reduction program.

1. Nice guy. I doubt my proclamation on the first day of class that I am a nice guy ever convinces anyone. However, in time they come to realize that I am doing my best to make statistics easy to understand and that I want them to succeed. I believe that probability and statistics, by giving us the tools to characterize risk and uncertainty and to make principled decisions in the face of uncertainty, are some of the most important intellectual discoveries in the last 500 years. Every class I try to share my excitement with the students. When I told a former student that I was writing this chapter, she said, "Be sure to tell them how you were always a cheerleader keeping our spirits up and convincing us we would eventually master the material if we only stayed with it." Still, although most students eventually consider me to be a nice guy, some remain puzzled as to why such a nice guy could be so excited about a topic like statistics.

2. Statistical software. There are some statistics instructors who believe that students only learn the material by doing hand calculations. I am in the other camp that argues that those hand calculations do more to teach arithmetic tricks using formulas that hide the fundamental statistical concepts. So I unabashedly allow the computers to do all the heavy lifting, leaving the students the responsibility of thinking conceptually about what the computer is doing for them. On the first day of class I show students how easy this is by having them open a dataset in a statistics program (we use StatView on Macs) and then calculate a mean for more than 1000 observations. When I first started using this approach many years ago, the students were often as afraid of the computer as they were of arithmetic, but that is no longer the case as few students are computer-phobic these days. Knowing that the computer will be available for the tedious arithmetic significantly reduces student anxiety.

3. Mastery exam. After covering the statistical material in the first half of the semester, students must pass a mastery exam to demonstrate that they have adequate statistical skills to do the projects in the second half of the course. Students must only pass this exam and they may have multiple tries if necessary to pass. Students' specific scores have no effect on their final grade so long as they pass the exam. The exam is in two parts: a closed-book test given during a lecture period and consisting of interpretations and critiques of output, statistical tests, newspaper articles, etc. and an open-book test given during the lab and consisting of datasets to be analyzed using statistical software and one small data problem for which the data must be entered and analyzed.

My experience suggests that students' anxiety is greatly reduced by not having the exam points determine the grade and by letting the students know that the possibility exists of retaking the exam if necessary, although few need to retake it. However, as we lead up to the mastery exam and I am helping students prepare for it, I have to be careful not to create the specter of the looming monster that will determine their fate. And it is a rigorous exam that thoroughly tests their skills. I use the analogy of a driver's license.

Passing the written and practical exam for a driver's license attests that the driver has sufficient skills to drive safely in normal conditions but does not guarantee that they could drive in adverse conditions nor that they know very much about what goes on under the hood. So too, passing the written and practical mastery exam in statistics attests that the student has sufficient skills to analyze data safely in the second half of the semester but does not guarantee that they could handle really messy data nor that they know very much about what is going on inside the statistical software we are using. The students all know that their driving skills have increased with practice and I assure them that their statistical skills will increase with practice on the projects later in the course. In any case, the metaphor of a driver's license appears to help them have realistic and less fearful expectations about the mastery exam.

4. StatFinder. Discussions with students during and after taking a statistics class reveal that they worry about finding the "right" statistical test. When doing homework problems in the t-test chapter, they do not have any difficulty realizing they need to be using t-tests. But when approaching a real data problem, students often fear that they do not know which statistical tests are appropriate. This is not surprising because most textbooks are organized with separate chapters for each test with little overall guidance about how to select tests, although increasing numbers of textbooks now provide a flowchart or some other aid for selecting tests. I provide students with a computerized flowchart called StatFinder: psych.colorado.edu/~mcclella/psych3101h/flowchart/statisticalProcedures.html.

My opinions about this and other such flowcharts are mixed. On the one hand, they necessarily make sharp distinctions among tests that are either very similar or sometimes identical (e.g., even though one can compute a t-test comparing two means by using regression with an indicator variable or contrast code, those tests are usually on distant branches of the flowchart tree) and they convey the incorrect impression that there is only one correct test for a given data question. Nevertheless, those disadvantages are outweighed for the beginning student, by giving them a good road map of where they might drive with their newly earned statistical driver's license. After they've had some driving experience I reveal that some of the tests on their road map aren't as far apart as they might first appear and that sometimes multiple destinations might be appropriate for the same goal. When I see a student sitting at the computer befuddled about what to do next, my first question is usually, "Have you tried StatFinder?" More often than not, that gets them going on the right road.

5. Pass/redo grading. All of the project activities in the second half of the course are graded "pass" or "redo." Although more than one "redo" is seldom necessary, students are relieved to know multiple tries are possible. Grades in the class are determined by the number and type of projects completed. This grading policy is based on sound cognitive psychology principles that emphasize repeated practice and time-on-task as important predictors of long-term mastery of the material. I once feared that a multiple-redo policy might encourage students to turn in shoddy work early just to see if it might pass. However, that has not been a problem. Students turning work in early are the good students and they do good work. The shoddy work is turned in at the very last minute by the students who procrastinated and are now hoping for the best. In any case, the grading policy changes the student strategy from "point-grubbing" to doing more work.

6. Codifying Commonsense. I try to convince students that much of what we do in statistics and in research design is nothing more than codifying commonsense notions they already have. Just as we would be suspicious if someone told us the score of a World Cup soccer match between two great teams was 27 to 13, statistical tests tell us when we should be suspicious of a surprising test statistic. And class discussion on the first day demonstrates students have quite a few commonsense notions about research design. I give students the discussion problem of how we might evaluate whether the unusual design of this course will help them learn statistics. The first suggestion is almost always to compare our section's scores at the end of the semester to those of another section using a more traditional method. But other students quickly raise the problem of so many uncontrollable confounding variables when there are only one or two other comparison sections. Recognition of the self-selection problem is more difficult, but usually one student will point that one out. Finally, it occurs to someone that we need a before and after design. Of course, I always have a test ready for that purpose! We take a quick-and-dirty twenty-item true/false test on the first day and we take the same test on the next-to-last day (so that students can do the statistical test evaluating the before-after difference on the last day). I think it important that we practice what we preach in statistics so that students can see the usefulness of statistical tools by analyzing data to solve problems of interest to them, such as this one.

The statistical content of my course is concentrated in the first half of the semester, and the focus is on data analysis. I use animated computer graphics (Java applets from *Seeing Statistics*, www.seeingstatistics.com) to provide illustrations of abstract concepts such as minimization of sums of squares, Central Limit Theorem, confidence interval coverage, etc.

I teach students that the underlying statistical ideas are the same for all common statistical procedures. Hence, we consider one in detail—the z-test for an hypothesized population mean—and use it as a model for all statistical tests considered later. I begin by using the students' scores from that twenty-item true/false test taken on the first day to ask the question of whether students were doing better than guessing? Did they know at least a little about statistics to start with? Typically, the mean score on the quiz is between 11.5 and 12.5. This provides a useful context for gently introducing many important statistical concepts. We discuss the following questions:

1. What mean score would we have expected if everyone had been guessing? *Null hypothesis is 10.*
2. Is a mean score of, say, 12 enough different from 10 to conclude that not everyone was guessing on all the questions? In other words, if everyone had been guessing, would 12 be a surprising mean score? This questions produces a variety of answers from the class and demonstrates the need for a principled answer.
3. Is a deviation of 2 from the expected value of 10 likely to occur by chance? We generate a first answer by having all the students take the test by guessing. To ensure they guess, they flip coins to determine their answers. For a class of 15 students, the 95% confidence interval for the null mean is between and 8.8 and 11.2; thus, it is not surprising that I've always been lucky with the coins' deviation being smaller than the actual class deviation of 2.
4. But can we depend on the deviation from one set of coin flips to be typical? The students have an intuitive notion that the one randomly-generated comparison deviation might have been "lucky" or "unlucky," and so that leads us to generating a more complete sampling distribution. An animated Java applet in *Seeing Statistics* allows them to generate an entire set of coin flips with a single button click. With multiple clicks they see (a) that the obtained mean class score of 12 is possible but very unusual (students shout out when they get a simulated score greater than 12) and (b) that the distribution of guessing means has the familiar "bell-shape" of the normal distribution. Students are then convinced that the class mean of 12 would be a very surprising value if everyone had been guessing.
5. Then, once they have these intuitive concepts we formalize it by discussing (a) statistical decision rules and Type I (false alarm) and Type II (miss) errors; and (b) the normal distribution, its approximation to the binomial, and z-scores. Then, if all goes well, they understand that the statistical result $z = (12 - 10)/.577 = 3.46$, $p = .0005$ implies that it would have been an extremely rare and surprising event to have obtained a class mean of 12 if everyone had been guessing so we ought to conclude that at least some students for some questions were not guessing.

After the detailed examination of the z-test for testing a population mean, it is then easy to present a template for all statistical tests we consider in the course:

a) Calculate the appropriate test statistic (e.g., z, t, F, r, R^2, χ^2) from the data.
b) Generate or find a table of the sampling distribution of the values of that test statistic we would expect to obtain "if nothing were going on." (Early in the course I use the vernacular "nothing going on" and only later move the students toward more formal language such as "null hypothesis.")
c) Compare the value of the test statistic to the sampling distribution to determine if the test statistic is "surprising."
d) If the test statistic is "surprising," then conclude that "something is going on." Otherwise, leave the question open.

With that structure, it is easy for us to move through a number of statistical tests (one-sample t-test, two-sample independent t-test, paired t-test, correlation and simple regression, multiple regression, one-way analysis of variance, repeated measures analysis of variance, two-way analysis of variance, and two-way chi-square independence tests). We let the computer program do the calculations and tell us what we need to know about the sampling distribution and then focus on interpreting the meaning of the statistical test and making sense of the output from the statistical software.

As we consider each new statistical test and its associated output, I emphasize that a complete analysis of any data question requires the answer to three related questions:

1. Is anything on? Answered with a null hypothesis significance test including its *p*-value.
2. How much is going on? Answered with effect size measures (although our software StatView is sometimes not up to the task) and confidence intervals.
3. What is going on? Answered by interpreting mean differences (and post-hoc tests if necessary) or percent differences or by interpreting regression coefficients.

To answer the final question about what is going on, students are required to describe their conclusions in substantive terms. This is especially important for a statistics course in a substantive discipline like psychology where the emphasis is on data analysis rather than statistics per se. I suggest to the students that they think about what they would say about the results on the evening news, where viewers won't be very interested in or have an understanding of statistical concepts and jargon. I also suggest that they re-read the last sentence in their summaries; if there is a number in that last sentence, then they probably didn't reach a succinct and appropriate conclusion.

With so much material to cover in the first half of the semester, there is not a lot of time devoted to issues in the "statistics and the citizen" category. However, I select a few such issues for special attention as they arise during the course. These are issues that I believe are important in the liberal education of any student. First is the issue of "compared to what?" Not specifying the comparison or confusing the comparison is a frequent problem that students are likely to encounter in media reports of statistical results.

A favorite example I use in class is an old newspaper article in which the reporter notes that 85% of shark victims are males and then speculates about why males must be tastier to sharks. Second is the issue of the two kinds of errors and their inherent tradeoff—focusing on one kind of error inherently increases the risk of the other kind of error. Public policy examples are easy to find.

As I write this, in the United States there are increasing complaints that there were enough warning signs to public agencies for them to have prevented the 9/11 terrorist acts (a Type II error), without considering what our lives would be like if the government acted on all warnings, resulting in many Type I errors. At the same time, there are increasing public concerns that the no tolerance policies for student threats and plastic knives in lunch sacks may be an "over-reaction" (Type I errors) as a response to the Columbine tragedy (a Type II error).

Discussing the two types of errors in policy contexts not only better informs students about the inherent difficulty of such decisions, but it also seems to help them understand the two kinds of errors in statistics and data analysis. Third is the issue of spread or variability around central tendencies. Being aware of the spread around expectations such as life expectancy, predicted weight, "normal" blood pressure, etc. can be quite liberating. However, I am less successful in communicating the importance of considering spread than I am in developing an understanding of the other two science-and-the-citizen topics.

A typical class

In the statistical content portion, roughly the first half of the semester, there are some similarities to a traditional class. I do spend a fair amount of time at the front of the classroom talking and writing on a white board. However, there are some important differences. I demonstrate everything on the computer—how to find assignments and syllabus on the web, how to use graphical animations in *Seeing Statistics* to understand statistical concepts, how to use statistical software, how to use StatFinder to find appropriate statistical tests, how to use library and web search engines—and then students *immediately* practice the same thing on their own computers. In a classroom of students and computers, for each demonstration some students will have problems either because they skipped a step as I was explaining it or because of the inevitable vagaries of computers and networks.

While the other students are exploring whatever was just demonstrated, I run—often literally—around the classroom helping those who are stuck for whatever reason. Particularly effective are demonstrations, either in the applets of *Seeing Statistics* or in the statistical analysis software, using our own class data. In the early class sessions, those data are the scores from the statistical knowledge quiz from the first day. We first examine it graphically, then we describe it numerically, then we simulate it, and finally we reach a formal statistical decision about whether they were guessing on the first day. And as we work our way through other statistical procedures, they hear me every day asking those three questions: Is anything going on? How much is going on? What is going on? I expect the students to answer those questions so as much as possible I use a discussion format. But when we get stuck, I am not hesitant to resort to an old-fashioned, but brief, lecture at the white board.

The first lab of the semester, held in the same classroom as the "lectures," is devoted to basic computer skills such as finding and opening the software package, using a browser to locate our class materials and the StatFinder, finding and using a word processing program, etc. As part of this lab, all students send me an email message introducing themselves. And I take digital photos of all students to create a photo roster for the class so that I and they can learn all the names quickly. The remaining six labs in the statistical content portion of the course are each devoted to one or two statistical techniques. A homework assignment is distributed (actually, it is made available in a folder on the course server) at the beginning of the lab.

Each homework set involves usually about half a dozen questions requiring the analysis of datasets (usually real) that are provided on the course server. Usually, one of the questions requires students to enter a small dataset for practice. Students are encouraged during the lab time to do the required analyses ("get the numbers") and then take them to a quiet place to think about them and to write their summaries. Although students have briefly practiced using the appropriate statistical software technique during class, a little review at the beginning of the lab is almost always helpful. Then I become a coach, moving around the lab helping those students who are stuck and staying out of the way of those students who don't need me. For the first few assignments the most common deficiency is a tendency to stop with the numbers of the statistical results. So I hector them frequently that they should not conclude with numbers but must instead conclude with a "5:00 News Summary" in which they explain the results in plain language.

The second half of the semester is devoted to practical application of the statistical tools learned in the first half of the semester. It would seem like it would be a good idea for each student (or perhaps small groups of students) to design a research project, collect the necessary data, analyze those data, and write a report. However, such projects have limited usefulness for several reasons. The data from such projects are often incomplete and messy and are therefore not appropriate for or beyond the scope of the statistical tools at the introductory level. Students learn more about the difficulties of collecting data than they do about the use of statistical techniques. A further special problem in psychology is that any research project involving human participants must be reviewed by the local Institutional Review Board with consequent delays that make completing a project by the end of the semester almost impossible. More importantly, what even experienced researchers learn from their first study on a topic is how they should design the next study differently. One-shot projects necessarily preclude the important learning experience of following up the results from one's previous studies.

Instead of one real research project, I decompose the process of doing a complete, multi-study research project into a series of separate activities and provide students with simulated experiments and previously-collected databases so that they can do multiple designs and analyses within one topic. Specifically, the following activities are assigned:

ExperSim: ExperSim originally referred to a collection of computer programs developed many years ago at the University of Michigan. These programs simulate experiments on several psychological topics (imprinting, social facilitation, etiology of schizophrenia, and motivation on routine tasks). An important feature of these simulations is that they give the student experimenter a number of possible design variables—some which have effects and some which do not—as well as levels of those variables—some of which are beyond the effective range of the variable. Several studies are usually required before a student has a refined experiment producing useful results. The computer returns results to the student experimenter almost as soon as the study is designed so there is ample opportunity to learn from one's design errors and from the data analysis to plan better follow-up studies. The ExperSim topics are unfortunately becoming somewhat dated but they still serve their purpose of providing students experience with the sequence of design-experiment-analyze-redesign. We have implemented a web interface to these simulations making them easy to use anywhere with an internet connection (visit samiam.colorado.edu/~mcclella/expersim/expersim.html).

The original ExperSim programs provide data appropriate for t-tests, analysis of variance, and two-way contingency tables, depending on topic. However, they seldom require the use of regression methods. I therefore provide some previously collected datasets that are more amenable to the use of regression techniques (e.g., college admissions and graduation data, course ratings, nationwide survey of mental health issues, etc.).

For an ExperSim project, students must complete a series of three related studies on one topic, where subsequent studies must build upon what was learned from earlier studies. Students write their project papers using the style of journal articles for the American Psychological Association. Papers in this style have introduction, method, results, and discussion sections for each study.

Library: In a lab session we learn about bibliographic database tools with special emphasis on those of use in psychological research such as PsychINFO and Social Sciences Citation Index (Web of Science). Students then select a topic (even perhaps one they need to explore for another class) and prepare a bibliography of 15 or more recent journal articles reporting empirical research on their topic. In addition, they write short reaction papers to three of the journal articles they read, with special attention to whether they could understand the statistics used. This library activity is meant to be an introduction to the kind of library research one does before beginning research in a new area.

Data Collection: What students would have learned about the difficulties of collecting data had they done their own research projects is important. To provide a bit of that experience we have a data collection activity in which each student collects a handful of observations as part of a larger research design. We use a protocol—usually from an ongoing research project being conducted by a researcher in our department—already approved by the local Institutional Review Board. The students then analyze the aggregated data and write a report describing the results of their analysis of those data and commenting on the difficulties of data collection they encountered such as problems of recruitment and random assignment, following experimental procedures uniformly, etc. Because they are untrained researchers and do experience unexpected problems when collecting the data, these data are only used by us internally and not for scientific publications.

A typical class in the practical application part of the course, roughly the second half of the semester, is very different from a traditional lecture. Most days in this part of the course are designated on the syllabus as "workshop." In the workshops, I do not lecture but instead work as a coach. I believe the goal in teaching is to make oneself obsolete. I am obsolete for some students sooner than others. Some are off and running on the projects and require no or very little help from me. Others, unfortunately dependent on having the teacher in classes always telling them exactly what to do next, have difficulty even selecting a topic, let alone knowing how to design a study to investigate that topic. Such students, even though they are by then quite comfortable with the statistical software, require extensive coaching. The greatest difficulty is a lack of critical thinking skills and this is seldom correlated with score on the mastery exam or past grades. Such skills are too seldom engaged in most classes. My coaching style is to ask leading questions to help them think through what the implications of alternative results would be for various hypotheses.

Often there is a group of students working on the same topic, so I pull them aside as a small group and get them to talk through the issues with me. I encourage them to investigate complementary research questions so that they can accumulate their knowledge and most importantly can help each other. I encourage collaboration and students do help each other a lot (critical feedback from a peer seems to sting less than from me), but each student is responsible for completing his or her own research project and write-up. The coaching style seems to work because towards the end of the semester I am indeed almost obsolete in the classroom as students busily work on their projects with almost no intervention by me.

In this practical portion of the class I do give a few lectures on topics announced on the syllabus. One of these lectures is how to write research papers in the style expected by the journals published by the American Psychological Association and explaining to students the reasons for the stylistic fussiness. In another lecture I describe the life cycle of real research papers from obtaining grant funds to submission to publication. Their **ExperSim** projects are meant to be a simulation of that process with me acting in the role of editor giving them either "accept" or "revise and resubmit" decisions on their manuscripts.

For one of their **ExperSim** projects, students give a short (10 to 15 minute) presentation to the class using a style one would use at a scientific meeting. The last four class sessions of the semester are reserved for this purpose.

The lab sessions in the practical portion of the class are devoted to individual activities. The early labs provide students with help on choosing a topic, using the experimental simulation software, and finding the databases for the projects. One lab is devoted to the library project. In this lab we learn to use electronic bibliographic databases provided by our library. We primarily use PsychINFO and the Social Science Citation Index from the Web of Science. With the increasing availability of electronic journal subscriptions through our library, students can often download the original articles they need for their library project. Another lab is devoted to the data collection activity. We practice the data collection procedures with each other, discuss issues of obtaining informed consent, and devise procedures for aggregating the data as they are collected.

To help students with their class presentations, one lab is devoted to the *proper* use of **PowerPoint**. As a psychologist, I am very particular that when preparing such presentations students must attend to such issues as readability (e.g., large, sans-serif fonts in colors having sharp contrast to plain background colors), visibility of graphics, and minimizing distractions (e.g., no cute transitions and fades with sounds). The final lab—if we are still talking to each other by that point in the semester and so far we always have been—is a celebration pasta dinner at my house.

Assessment

To pass the course, students must pass the mastery exam and participate in the evaluation activities (completing the university's course evaluation form and retaking the same statistical knowledge quiz we did on the first day of class) at the end of the semester. But that only would earn a "D." To earn a "C," students must in addition complete one of the activities in the second half of the course. To earn a "B," students must complete an **ExperSim** project consisting of three related experiments or studies and one of the other activities. To earn an "A," students must do two **ExperSim** projects, each consisting of at least three related experiments or studies, and they must complete both the **Library** and **Data Collection** activities.

In terms of evaluating the effectiveness of the course, I begin by using as a class discussion topic how we could assess whether the unusual structure of my course provides educational benefits. The students are quite adept by the end of the semester in pointing out the difficulties of making comparisons across sections including non-independence of individual student data within sections and self-selection into sections on the basis of time and published instructor teaching ratings.

We can and do evaluate whether students improved during the semester. On the next to last day of class (so that we can analyze the data on the last day), students retake the same twenty-item statistical knowledge quiz they took on the first day of class. The test has no effect on their course grade and no student has ever reported studying for it. And for the last half of the semester they have been working on projects rather far removed from the attention to textbook definitions in the first half of the course. Nevertheless, the paired *t*-test comparing end-of-term and beginning-of-term scores is always about 4, reflecting an improvement of about 1.5 standard deviations.

Given that educational interventions are usually considered to be successful if they achieve an improvement of half of a standard deviation, I am of course quite pleased by that assessment. However, as my students are quick to point out, it does not rule out the possibility that these or similar students might have improved even more with a more traditional course. Furthermore, the twenty-item true/false test necessarily tests knowledge of statistical concepts and does not assess achievement of my higher-level goal of teaching them how to use statistics in the context of conducting thoughtful psychological research. The assessment of that is informal and comes from the instructors who teach the upper division lab courses. They report that my students are the ones who still remember from their statistics course how to do the analyses and equally important how to report them in the style of American Psychological Association journals. They also report my students are frequently observed helping the other students who are struggling.

An even more gratifying informal assessment happened last semester. Two students were using data for one of their ExperSim projects that they had obtained from their respective labs where they were working as undergraduate research assistants. I noted to myself that the questions they wanted to ask of their datasets would require the use of statistical techniques (mixed-model analysis of variance) that we had not covered in class and that StatFinder only identified as an "advanced topic" with no details. However, our statistical software could do the analysis. I decided to allow these students to struggle with their respective problems for a while and then I would intervene to explain this "advanced topic" to them. I was surprised—and extremely pleased—when they called me over not to ask for help with the analysis that they had already correctly done, but only to check their substantive interpretation of their statistical results. If the course got these students to a level where they could successfully go beyond the material we had covered, then I was truly obsolete and the course was a success.

Developing an innovative course

After my efforts to teach a theoretical course failed so miserably in my first year of teaching, I had to find a different approach. I naturally approached the problem from a psychological perspective and analyzed the statistics course to identify the obstacles that prevented students from learning. I then designed my course to accomplish three goals: reduce fear, practice relevant skills, and increase time practicing those skills.

1. Fear. Fearful students simply cannot learn the material; instead they learn defensive skills just to survive the course. My strategies and tactics for reducing fear are described in detail above. I remain convinced this is the single most important issue in my course. If fear isn't reduced, then all other teaching efforts are wasted.

2. Practice Relevant Skills. Many research studies in cognitive psychology have demonstrated that it is difficult to generalize concepts and knowledge, especially newly acquired abstract knowledge, to new domains. For example, students taught the Central Limit Theorem will be unlikely to see spontaneously its application in contexts other than the exact one in which it was taught. The inference then is that if I wanted students to be skillful at doing psychological research then I needed to have them practice the specific skills they would need. Those skills included data analysis but a lot of other things as well. The skills I want them to have are reflected by the activities in the second half of the course as described above. To make room for practice on these skills inevitably meant much less emphasis on the memorization of definitions and formulas that often characterize more traditional courses. And it also meant more of a cookbook approach to learning the statistical tools themselves. Also, I wanted students to be able to write cogently about their data and what it meant. And I didn't particularly care if they were good at taking multiple choice tests. Hence, *every* homework assignment, *every* question on the mastery exam, and *every* activity in the second half of the course involves writing.

3. Increase Time on Task. A permanent issue in psychology is how much of our behavior or performance is due to nature versus nurture or to talent versus practice. But everyone agrees that whatever one's nature or talent, skills improve with practice. Diary studies reveal that expertise at the highest levels is ultimately determined by practice—concert soloists practice more and top athletes train more. Hence, I wanted to increase the time students spent practicing those skills, especially analyzing and interpreting data, that I thought were important for conducting psychological research. I therefore changed the usual grading structure that rewards quality on a fixed quantity of assignments and tests to one that rewarded quantity so long as the quality was minimally acceptable. Not only is this better for student learning but it also greatly improved my relationship with the students. Instead of being the ogre telling them that the points they lost on the exam could never be made up, I now provide constructive feedback on how to improve their work when they redo it.

There were relatively few obstacles to developing an innovative course in my department other than the usual conflicting demands on one's time at a major research university where grants and publications are expected. Even though my department offers three to four sections of the introductory statistics course each semester, there is little coordination of content and instructors are free to use any textbook and to approach the course in whatever manner they choose. But neither are there rewards for developing an innovative course other than my personal satisfaction that I am doing a better job and the occasional comment from other professors when they've noted the students who can do good data analyses have come from my class. The biggest reward is that I have so much fun teaching this innovative course; after 27 years of teaching I remain far from being burned out.

When I first developed this course I imagined obstacles that were not there. I worried that I might not have correctly calibrated the amount of work appropriate for the different grades so that my grade distribution would be sufficiently anomalous to warrant a call from my department chair or even the dean. I was particularly worried that the quantity-based grading structure would make a high grade within reach of every student so that the class average would be extremely high. However, if the students did all that work then I would have no qualms about giving a lot of high grades. To protect myself when the dean called, I kept all the student papers so that I could demonstrate the quantity and quality of the work that justified those high grades. Enough students opted for the quantity of work that gave them an "acceptable" grade so that my class average, although higher, was not outside the range that would generate any notice. Having become a better cheerleader and teaching mostly honors students, my grade distribution has steadily moved upward but the dean has yet to call.

The only important obstacle with my course's increased emphasis on doing data analysis has been the same one I first faced in high school—adequate computational tools. In the early years I required students to have decent pocket calculators and used Linton and Gallo's *The Practical Statistician*, an unabashed statistical cookbook that gave step-by-step instructions (with lots of internal checks) on how to perform a wide array of statistical calculations using pocket calculators. **ExperSim** gave students a few basic statistics such as means, standard deviations, and sums of squares from which it was easy to calculate test statistics using the *Practical Statistician* recipes. Many copies of that book disappeared from my office as it was an instant hit not only with undergraduates doing theses but also with many graduate students.

Although this approach was very popular with the students (so long as they didn't have to do too many analyses) because it empowered them, it was still far too primitive to allow either big datasets or multiple regression analyses. Over the years the computing resources have steadily improved to the point that I now teach in a classroom with the latest flat-screen, minimal-footprint personal computers for every student. However, there always seem to be battles with those who administer the computer labs. Their basic software load set for laboratory computers is oriented towards word processing and web browsing so one needs to be vigilant to make sure they don't forget the statistical software. And even then the statistical software is not widely available so students in my class are constrained to do their work in relatively few of the many campus computing laboratories. The computing staff also seem to work on a schedule that is independent of the academic calendar—they have a penchant for making major changes in mid-October just after all the computing issues have settled down in class. Nevertheless, when I remember using adding machines and slide rules or carrying large boxes of punched cards across campus, I am extremely grateful for the wonderful computing tools available to me and my students.

Teaching in a computer lab has produced one humorous problem involving social norms that I find fascinating as a psychologist but extremely irritating as an instructor. The computers and monitors generate enough heat that we usually need to leave the door open. Although students would never think of entering a regular classroom during a lecture to find an open desk at which they could write a letter, they seem to think it entirely appropriate to enter a computer classroom during a lecture to find an empty computer at which they can send email and browse the web. When I am explaining some difficult concept, some students just march in and others think they are being polite by interrupting and asking if it would be OK to do email while I lectured if they promise to be quiet. My strong reaction to those incidents has earned me among students not in my class the nickname "Scary Gary." Nevertheless, more than once I've had to call the campus police to remove people who were working in the lab before class and then refused to yield the computer to a student in my class.

After that initial failure with the theory-based statistics course, the basic structure of my present course has been in place for many years. But many of the specific details have evolved over the years, mostly in response to new opportunities. For example, for the library lab we used to troop over to the library where a librarian taught the students how to use the printed bibliographic reference tools. Now we stay in our regular computer lab and do all our searching and even much of our journal reading electronically. There have also been innovations that I soon discarded as bad ideas. It is worth considering one of those bad ideas in detail.

I once read a definition of the lecture as "the mysterious process by which the notes in the professor's notebook are transmitted to the student's notebook without having gone through the head of either one." This process is much less mysterious these days as it has become institutionalized with many professors making their lecture notes available as *PowerPoint* slides that students may download and print.

Although I should have known better, I thought that perhaps the rapid pace of the first half of the semester might be more tractable if I made more of the material, especially the equations, available as text that they could follow along as I talked. Making the slides was a lot of work so I was disappointed when they didn't seem to be helping; in fact, I thought they made the class lifeless. I stopped in the middle of class one day and asked the students (remember these were honors students) if the slides were a good idea. The nearly unanimous opinion was that they were a bad idea because the students just leaned back and didn't have to work to get the information. I happily stopped doing the slides and ever since have been vigilant to engage the students with as many "lean-forward" activities as possible and to avoid "lean-back" instruction at all costs. Hence, even during my lectures there are frequent breaks for the students to do on their computers whatever it is we are discussing. And I also avoid videos, animations, and other multi-media that might induce "lean-back" learning.

Part of developing my course involved creating *Seeing Statistics*, an online statistics textbook published by Duxbury Press (www.seeingstatistics.com). An online textbook is not an essential ingredient for this course as I have previously taught it successfully using a more traditional printed textbook. However, given that all other aspects of the course are computerized, it makes more sense to have the textbook computerized as well. I was first motivated to write *Seeing Statistics* by students asking, "Can't we have more pictures?" Probability and statistics are inherently geometric with heavy use of geometric concepts such as areas under curves, sums of squares, etc. Yet most traditional textbooks teach statistics algebraically instead of geometrically. The pictures I could produce by drawing on the white board or by using the simple graphing tools in word processing programs

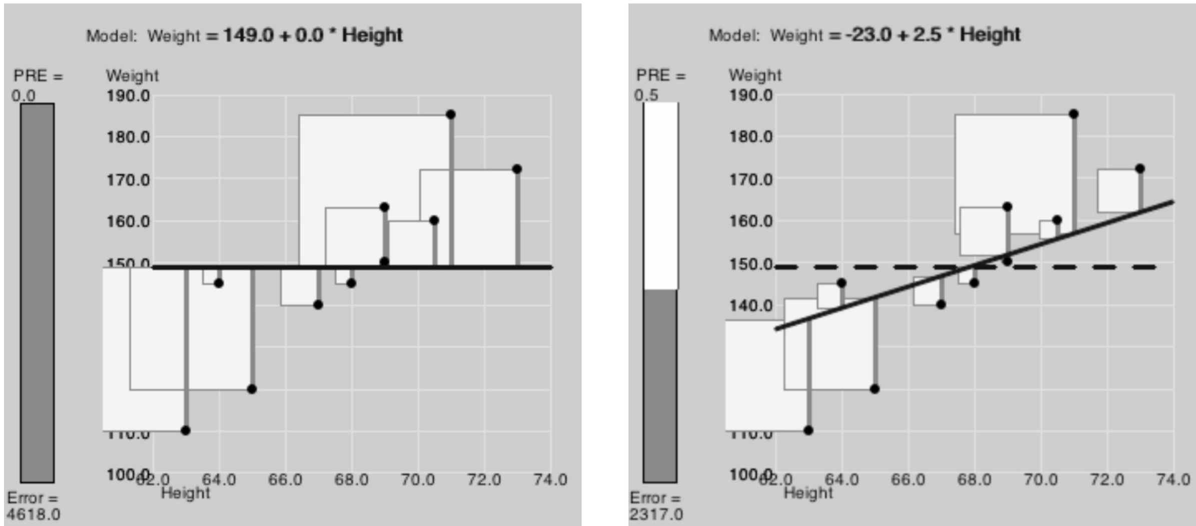


Figure 1. Two views of a Java applet from *Seeing Statistics* illustrating fitting a regression line by minimizing the sum of squared errors.

were always quite limited and necessarily static. Tools available to online authors, such as Java applets, make it possible for even those of us who are artistically challenged to make graphics that are not only more interesting but more importantly are dynamic. Being able to manipulate the graphical images helps students understand the statistical concepts. And the manipulation is very important because it converts the lean-back learning of automated animations into lean-forward learning.

Figure 1 depicts one example of finding the best-fitting regression line by minimizing the sum of squared error. The first panel shows the data fit with a line through the mean and the resulting squared errors. The meter on the left shows the total sum of squared errors. The second panel shows the line after it has been moved by the student so as to reduce the sum of squared errors. The student can see (much more easily when this is dynamic on the website than in these two static images) the squares becoming smaller and the total error getting smaller in the meter on the left (the proportion of the error reduced—equivalent to r^2—is represented by the light portion of the error meter). The second panel still requires further manipulation to find the minimum sum of squares and the best-fitting regression line.

Figure 2 depicts one image from an applet representing a scatterplot and its associated correlation. Most publishers of printed textbooks will allow authors to include only three or four different scatterplots to illustrate the relationship between the organization of the points and the correlation coefficient. In contrast, the student using the online applet depicted in Figure 2 can move the slider bar to see hundreds of different scatterplots and their correlation coefficients, with each plot smoothly transforming into the next. The applet, running on the student's computer, constructs each image as needed so it is much more efficient than downloading multiple GIF or JPEG images of those scatterplots.

Although a few students report that they would prefer to have a printed textbook to keep and to write in its margins, most students report preferring the online textbook because it is cheaper (without the cost of the paper and printing) and often handier. Also, the same applets used in *Seeing Statistics* for illustrating statistical concepts work well as interactive displays when projected during

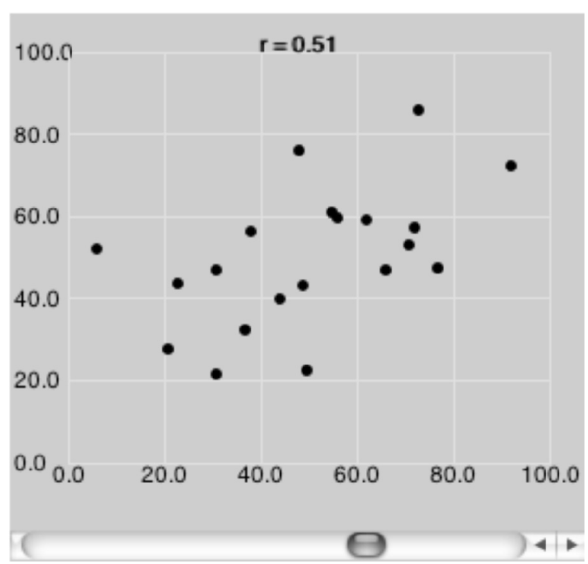


Figure 2. Java applet from *Seeing Statistics* illustrating the relationship between the scatterplot and the correlation coefficient with slider for changing the degree of relationship.

the lecture. For many of the applets, it is easy to paste in data from the class, making the illustrations of more interest to the class. Readers of this chapter interested in trying out *Seeing Statistics* should contact the author at gary.mcclelland@colorado.edu.

Future plans

The basic structure of my class has been both successful for the students and fun for me to teach so I doubt I will change it in the dwindling years of my teaching lifetime. I will however continue to tinker with the details, but it is difficult to predict what those will be because they depend on the opportunities that new technologies and tools will present. Much of my future work will likely be directed toward further development of *Seeing Statistics.*

We have much to learn about using interactive graphics in education and not only my teaching but also much of my scientific research will be devoted to understanding how to use interactive graphics effectively. To become professors we demonstrated verbal skills by writing a dissertation and defending it orally. And to guide us we had style manuals and a grammar developed over hundreds of years. Few, if any, of us were selected because of our ability to use graphical representations effectively. The few graphical style manuals available were motivated more to make it easier to draw static graphs with mechanical draftsman tools than by a desire to communicate information effectively. And we are just beginning to start constructing the equivalent of a grammar for graphical objects. Interactivity adds even more opportunities and choices for which we do not yet know the best methods.

I was not long at work on *Seeing Statistics* before I realized there would be accessibility problems. Subsequently I had the opportunity to teach statistics to two blind students, one of whom taught me a lot about screen readers, tactile devices, and other such tools. Together we experimented with various ways of listening to data and discovered that time trends and outliers are often more noticeable aurally than visually. My sequel will be *Hearing Statistics.* Not only will this ameliorate the accessibility problems inherent with *Seeing Statistics* but also, just as curb cuts are good for everyone, I believe that auditory representations of data will be useful for everyone. Using special heat-sensitive paper that rises where there are photocopied lines, we felt graphs of theoretical distributions and data. The more limited resolution of touch and the difficulty of maintaining orientation made our feeling experiments much less successful. Nevertheless, as tactile devices improve I am optimistic I'll eventually produce *Feeling Statistics*. My personal view as a psychologist is that the available research on different learning styles has been greatly overstated. However, using more sense modalities than just vision to understand data should be useful to everyone. Nevertheless, I have no immediate plans ever to develop *Tasting and Smelling Statistics*, but who knows what tomorrow's technologies will allow us to do.

The extensive writing requirements necessitate a lot of reading and feedback from me. This would make it difficult to scale the course to much larger class sizes. My colleague Tom Landauer and his associates are making great progress with the computer assessment of written materials. I plan soon to experiment with these techniques to provide feedback to students about their written summaries of data. If successful, this would allow even more practice on the important skill of interpreting and summarizing data.

If I knew what to do, I would change my course so as to increase student motivation. Although I sometimes fear that they are simply the observations of an ageing professor who believes things were better in the old days, I do believe that today's students are less motivated. Many seem to be going to college only because that is the next thing one does after high school to avoid entering the real work world a little longer. If I knew how to get students to be passionate about statistics or any other intellectual topic, I'd do it. My discipline of psychology has much more guidance to offer about cognitive issues than motivational problems. However, more and more psychologists are turning to the study of motivational issues and I'll be monitoring that research for whatever insights might be useful in teaching statistics to psychology students.

In the meantime, I am eagerly awaiting my next class session in which I have the privilege of teaching statistics to psychology students.

15

Teaching Statistics to Business Students

Jonathan D. Cryer
University of Iowa

I have found it very helpful to be active in organizations on the forefront of statistics education.

Becoming a teacher of statistics

In 1957 I enrolled in a fine liberal arts school, DePauw University in Greencastle, Indiana. My intent was to become an engineer through their 3-2 program—three years at DePauw, and then two years at an engineering school affiliated with this program. However, after getting into my first real mathematics courses at DePauw, I decided that mathematics was my real love. In my junior year I took my first statistics course, really mathematical statistics, using Paul Hoel's book. It was taught by Robert Hultquist, who had just completed his PhD in statistics at Oklahoma State University. I was hooked! Statistics looked like an area where I could use mathematics and science (I minored in physics) to help solve problems of a substantive nature.

In 1961 I was fortunate to be awarded a National Defense Education Act Fellowship to continue my study of statistics in graduate school at the University of North Carolina–Chapel Hill. During my final two years in graduate school I was hired by Gertrude Cox to work at the Research Triangle Institute (RTI) near Chapel Hill. There I worked on several projects and saw how statistics could impact the real world. During this period I taught briefly at UNC when my major professor needed his course covered during his extended stay abroad. I also taught a graduate course in time series (never having taken such a course!) at North Carolina State University, taking over for a professor who had taken ill. With support from RTI and NASA projects I received my PhD in statistics from the University of North Carolina in 1966 and subsequently began teaching full time in the Department of Statistics at the University of Iowa.

In the subsequent 35 years I have taught many different courses: a freshman seminar on the graphical display of quantitative information, business statistics for both undergraduates and MBA students, mathematical statistics for both undergraduates and PhD statistics majors, and regression and time series methods.

One of our bread-and-butter courses is Statistics for Business. It enrolls lots of students and supports the vast majority of graduate students through teaching assistantships. As I was not happy with the content and direction of this course I became involved with the conference, *Making Statistics More Effective in Schools and Business* in 1988 when the conference was held at the University of Wisconsin. This series of conferences began in 1986 under the leadership of Professors Harry Roberts and George Tiao, both of the Graduate School of Business, The University of Chicago. The mission of these conferences was to improve the teaching and practice of statistics in Schools of Business. The aim was to encourage interaction between business faculty and others involved in teaching business statistics to business students, as well as interaction with professionals from industry and government, with publishers, and with software producers. Information about future and past conferences may be found at their Web site (www.msmesb.org).

I have participated in most of the conferences, served on the planning committee, and hosted the conference at the University of Iowa in 1996. The interaction with colleagues at this conference and elsewhere has been invaluable in helping me develop as a teacher and putting together a modern, innovative business statistics course.

My course

The University of Iowa is a state supported university consisting of many colleges. The total enrollment is about 29,000 students with the majority (15,000) enrolled in the College of Liberal Arts and Sciences. I am a member of the Department of Statistics and Actuarial Science which is housed within the Division of Mathematical Sciences in the College of Liberal Arts and Sciences.

The course, *Statistics for Business* (22S:008), is taken mostly by pre-business majors who are presently enrolled in the College of Liberal Arts and Sciences but who hope to be admitted to the College of Business after completing this course and other business college prerequisites. The students are predominantly freshman and sophomores with a smattering of juniors and seniors. This four-semester-hour course is taught every semester and enrolls about 1200 students per year. During the fall and spring semesters the course is taught in a very large lecture accompanied by small discussion sections. Students attend three lectures and two discussion sections each week.

Every week one of the two discussion sections meets in a computer laboratory classroom. The other day discussion sections meet in a regular classroom. Faculty deliver all of the lectures while graduate student Teaching Assistants (TAs) handle all of the discussion sections with the overall supervision of the faculty member. We use the textbook, *Statistics for Business: Data Analysis and Modeling*, (Cryer and Miller, second edition, 1994), designed to meet the goals of a modern course in statistics for business students. Much of the impetus for writing this book came from attending the conferences *Making Statistics More Effective in Schools and Business.*

Statistics is the systematic study of variation in data: how to measure it, display it, summarize it, model it, and use it to learn new knowledge. The goals of this course are for students to learn about and do all of these aspects of statistics and to use them in effective and appropriate ways. This is neither a mathematics course nor a "baby" mathematical statistics course. Mathematics (namely algebra) is used occasionally but only when it helps in learning statistics. This also means that probability is not emphasized. Probability is introduced as needed along the way with the emphasis on interpretation as a long-run relative frequency. I have found that probability "rules" are rarely needed if probability is introduced intuitively and analogies to finding areas are given.

This course is one of the largest classes at the University of Iowa. The three weekly lectures are presented in a large lecture hall equipped with a rear projection system and a very large screen. The computer screen, videos, and handwritten material may all be easily projected on the large screen. Lighting is controlled from the lecturer's console and can be set at a reasonably bright level without obscuring the material on the large screen. It is usually dimmed a bit only when showing videos such as the *Statistics: Decisions Through Data* series. "Real time" computing is used frequently to illustrate many of the concepts. This is especially useful for doing simulations and "what if" scenarios.

Although the lectures are very large, (550–600 students), I always attempt to move around the room, talk directly to individuals, and get their input on the matters currently under discussion. I use a wireless microphone PA system and find this a real advantage. The lecture notes are available to students (and everyone else worldwide) on the internet both before and after class. See www.stat.uiowa.edu/~jcryer/22s008.html. In general, the notes do not contain complete detailed lectures but do help minimize the hassle of frantic note taking. (On the minus side, some students feel it is not necessary to attend lectures since the notes are always available online.) I usually display the notes on the large screen to guide the day's lecture. Since there is only one large screen, I switch back and forth between Minitab displays, lecture note displays, and handwritten displays as needed and appropriate.

Topics covered

I cover many standard topics but also include numerous important nonstandard topics. The following detailed list shows the difference between topics in this course and a typical introductory statistics course:

- Problem formulation tools: flow diagrams, operational definitions, cause-and-effect analysis, Pareto analysis
- Plotting process data
- Plotting distributions
- Summarizing continuous data
- Describing categorical variables
- Relating continuous variables
- Multiple regression models
- Normal distributions
- Control charts for continuous variables
- Binomial distributions with an introduction to hypothesis testing
- Control charts for binary variables
- Data collection: surveys and experiments
- Sampling and confidence intervals for proportions
- Significance tests, confidence intervals, and prediction intervals for means

The discussion sections are used to answer student questions, discuss lecture topics further, discuss homework questions and their solutions, carry out activities in small groups, and so forth. The weekly day in the computer lab is used to help students with the Minitab software that is used throughout the course. Some days the lab session helps with routine use of Minitab to solve typical statistics problems. Other days we do substantial simulations to illustrate statistical theory. I have written numerous interactive Minitab macros that facilitate graphical illustration of confidence intervals and their meaning, graphical display of areas under normal curves, robustness of the confidence interval based on the t distribution, lack of robustness of confidence intervals using the z statistic, and so on. These macros are used both in Lab and lecture. I have also written a *Minitab Handbook* of over 200 pages designed specifically for this course. The *Minitab Handbook* is custom published so that it may be updated each year as the software changes. Assignments which require computer work are given nearly every week. These assignments are seldom simple calculations of statistics—rather they often require manipulating data to look at what happens with and without outliers, producing graphs with males and females combined and separated, producing graphs with and without a categorical variable determining the plotting symbol, and so forth. Similar exercises are solved and illustrated in the Lab but students are expected to work on the assignments outside of class time. They must be written up in a document with actual sentences explaining what is done and what it means. Courses which do not use good statistical software or use statistical calculators cannot carry out these activities without overburdening the student.

A typical class

I do not believe that there is such a thing as a typical class, so I will describe an example of a class. This example is a class where we are first studying correlation. I begin by setting out the topics to be covered in this unit. The topics listed are:

- Relating Continuous Variables
- Scatterplots and Relationships between Variables
- Positive and Negative Associations
- The Correlation Coefficient
- Limitations of Correlation
- Causation

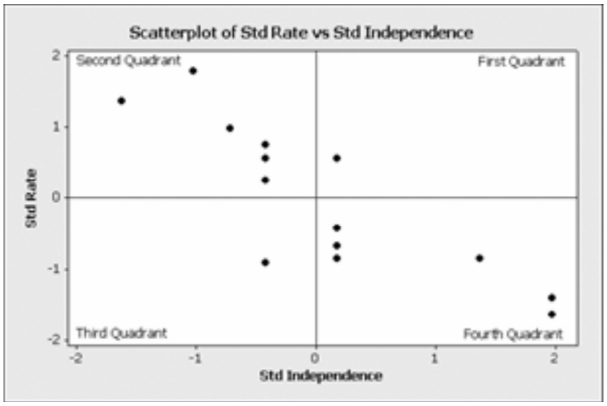
I then use the in-class computer and Minitab software to display a scatterplot of the following data:

Inflation Rates and Central Bank Independence, 1955–1988

Rate	Independence	Country
7.8	1.00	New Zealand
8.5	1.50	Spain
7.2	1.75	Italy
6.8	2.00	United Kingdom
6.5	2.00	Australia
6.0	2.00	France
6.0	2.00	Norway
6.0	2.00	Sweden
4.1	2.00	Belgium
6.5	2.50	Denmark
4.9	2.50	Japan
4.5	2.50	Canada
4.2	2.50	The Netherlands
4.2	3.50	United States
3.3	4.00	Switzerland
2.9	4.00	Germany

The scatterplot of these data shows a clear decreasing association—as degree of central bank independence increases, inflation rate tends to decrease. I suggest that it would be nice to have a numerical measure of the magnitude and direction of the association between these two variables. This leads to the verbal definition of the correlation coefficient. I then explain in words the process of computing the correlation coefficient from standardized values, followed by a formula with symbols for those students who appreciate such things. I always try to give both a verbal step-by-step explanation and a formula since some students relate better to one or the other of these formats.

I find that a scatterplot of standardized values is helpful in understanding why the correlation coefficient is positive for increasing associations and negative for decreasing associations. Here is such a plot for the inflation rate data:



From such a plot it is easy to see that products coming from the first and third quadrants will give positive contributions to r and products from the second and fourth quadrants will contribute negative terms to the correlation coefficient r. A predominance of pairs in the second and fourth quadrants leads to a negative correlation coefficient in this case.

We discuss the usual facts about correlation and examine different sets of data where there are different patterns or relationships but each one has a correlation coefficient of close to 0. We discuss why this value is obtained and what it tells and does not tell about the data set. I think it is important to have the students see many examples of bivariate data sets, with plots that illustrate that a curved relationship can have a strong positive or negative correlation coefficient even though the data do not follow a straight-line trend. This also helps set the foundations for the next unit where we model curved relationships.

Using the Minitab software we discover the limitations of the correlation coefficient—the effect of outliers, lurking variables, and the difficulty of establishing cause-and-effect. The effect of outliers on the correlation coefficient is easy to demonstrate by putting up nearly any scatterplot and asking what would happen if a particular data point were changed. The software allows us to easily reconstruct the graph and recompute the statistics.

I always include a discussion of lurking variables, where the relationship between x and y may be caused, wholly or in part, by a third variable z that is related to both x and y. We discuss examples of lurking variables such as:

- There is a high positive correlation between the number of fire trucks on the scene and the amount of damage suffered.
- A study of hospitals finds a high positive correlation between the number of medical specialists on staff and the death rates of patients.

I challenge my students to determine what the lurking variable is in these examples. (This last example is especially easy to relate to in Iowa City where University Hospital has lots of medical specialists.)

For practice in seeing and guessing correlation I use two types of activities. I have a Minitab macro which lets the user specify the correlation coefficient, then the software generates the data, plots it, and shows the specified correlation. This helps students see different types of scatterplots that match specific correlation coefficients. We also use the Guessing Correlation Web Applet (see Robin Lock's chapter) to help students correctly match correlation coefficients to the correct scatterplot. (However, I do think that game has some serious limitations. I do not think that even the best statisticians can differentiate between a plot with $r = 0.862$ and one with $r = 0.864$ nor is it useful to do so.)

Assessment

Assessment in this class is rather traditional. Quizzes are given weekly, sometimes during lecture and sometimes in discussion sections. Quiz times and places are not announced ahead of time. This gives some added incentive for the students to attend the lecture. Several quizzes involve computing with Minitab. These are given in the Computing Lab. Homework is assigned and collected and graded by the TAs nearly every week.

Common midterm examinations are given two or three times throughout the semester and a common final exam is administered during finals week. All sections are graded on a common, fixed scale that is announced on the first day of class in the syllabus. See www.stat.uiowa.edu/~jcryer/22s008.html for the complete syllabus and copies of many examinations from past years.

For a number of years students were assigned Team Projects. Teams consisted of three or four students who designed their project from scratch with guidance from their TAs and a four page handout from the instructor. Students were expected to investigate an area of interest to them from initial description of interesting questions, data collection issues, the data collection process, through analysis of the data, and most importantly, the final written report on what was learned. Although I believe these projects were fairly effective, they were quite difficult to administer in a large class. The TAs have had very little experience with projects themselves and the faculty member must expend a great deal of effort guiding the assessment of the project proposals and the final write-up. We are not currently assigning projects in this course but they are optional assessment items in the required follow-up course offered in the College of Business.

Developing an innovative statistics course

Developing an innovative statistics course has required a substantial time commitment. Changes to a traditional lecture class take time and do not always receive support from the university and department. I have found it very important to be active in organizations on the forefront of statistics education, such as The American Statistical Association and the *Making Statistics More Effective in Schools and Business* conferences.

An innovative course requires innovative teaching materials and adequate computing resources: new textbooks, software manuals, different types of examinations and other assessment are needed. With regard to textbooks, Bob Miller (University of Wisconsin) and I took on the daunting task of writing an appropriate textbook for a modern course in statistics for business.

Adequate computing resources are a must for modern statistics. I cannot imagine teaching statistics without such resources. Even mathematical statistics can benefit from appropriate use of the computer. We are quite fortunate at the University of Iowa to have excellent computing facilities. In particular, our Department was the recipient of a substantial gift from an alumnus, Robert Myers, which we used to equip our Computer Lab—the Robert and Rudy Myers Computing Laboratory. We use this room continuously from 7:30 a.m. until 6:30 p.m. on Tuesdays and Thursdays for *Statistics for Business* discussion sections. The University has about 25 such labs across campus that are open many hours per week. Students may use Minitab and access the course datasets from any of them. Of course, a great many students have their own computers in their residences. The Minitab software may be purchased or rented at a very reasonable cost for these machines.

Future plans

The future of our course is unclear due to my retirement in 2001. It will be up to the new and younger faculty to chart the course for the future.

References

Cryer, J. D. (2003). *Minitab Handbook to Accompany Statistics for Business, Data Analysis and Modeling.* (6th ed.) Plymouth, Michigan: Hayden-McNeil Publishers.

Cryer, J. D., web page for Statistics for Business, 22S:008 Available at www.stat.uiowa.edu/~jcryer/22s008.html (Or click link on my home page at www.stat.uiowa.edu/~jcryer)

Cryer, J. D. and Miller, R. B. (1994). *Statistics for Business: Data Analysis and Modeling* (2nd ed.). Belmont, CA: Duxbury Press, an imprint of Wadsworth Publishing Company.

Statistics: Decisions Through Data video series. Reading, Massachusetts: Addison-Wesley Publishing.

About the Editor

Joan Garfield is a Professor Educational Psychology at the University of Minnesota where she directs a graduate program in Statistics Education. She is the author of many articles on teaching and learning statistics and has co-edited three previous books on this subject. She helped to establish CAUSE (The Consortium for the Advancement of Undergraduate Statistics Education) and serves as Associate Director for Research. She co-chaired the GAISE project (Guidelines for Assessment and Instruction in Statistics Education). She is active in the national and international statistics education communities and has served as Vice-President of IASE and Chair of the ASA Section on Statistics Education. An award-winning teacher of statistics, she now directs her efforts towards courses and programs preparing future teachers of statistics.